CURING OBESITY • RING BUBBLES OF DOLPHINS • THE SUN'S MAGNETIC FIELD

SCIENTIFIC AMERICAN

AUGUST 1996 $4.95 U.K. £3.00

THE ELECTRONIC WALLET:

NEW "SMART" CARDS COULD REPLACE CASH, CARRY DATA AND IMPROVE TELECOMMUNICATIONS

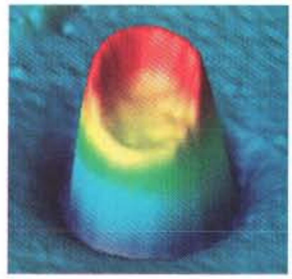
Look closely at sand and see the earth's history

08>

9 770036 873039

SCIENTIFIC AMERICAN

August 1996 Volume 275 Number 2



REVIEWS AND COMMENTARIES

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 1996 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652; QST No. Q1015332537. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send e-mail to info@sciam.com Visit our World Wide Web site at http://www.sciam.com/ **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

About the Cover

The myriad shapes of Japanese star sand (magnified eight times) reveal the grains' diverse origins as shells and bits of stone. Image by Laurie Grace, from photography by Christopher Burke, Quesada/Burke Studios.

FROM THE EDITORS

Aliens at Play

The classic bugaboo of animal behavior research is the sin of anthropomorphism: Thou Shalt Not Think of the Beast as Man. No matter how much an animal may seem to act like a person, professors sternly warn students, never forget that millions of years of evolution mentally separate the two. I once made the mistake of smiling at a cute rhesus monkey—forgetting that among its kind, bared teeth are a call to battle. Ever seen the incisors on a rhesus monkey? They're *sharp*.

Seeing ourselves in animals, and animals in ourselves, seems inescapable. We cannot scientifically quantify our emotional kinship, but we cannot disregard it either. Pet owners vouch for the capacity of cats, dogs and other creatures to be proud, lonely, disdainful, embarrassed and more. Meanwhile we laugh like hyenas. We preen like peacocks. We show the courage of lions and the cunning of wolves and the bland obedience of sheep.

ROBERTO OSTI

DOLPHIN FUN sometimes involves playing with hoops made of air.

Sometimes, though, animal studies afford a chance to feel at once the similarity and the strangeness of nonhuman minds. Consider the glimpse of dolphins that Ken Marten and his colleagues offer in "Ring Bubbles of Dolphins," on page 64. Television and movies portray the cetacean star Flipper as a loyal, dependable pet who loves human company—Lassie with a blowhole. (And Lassie, very clearly, is a Boy Scout in a dog suit.) But that comparison does dolphins a disservice.

These are shrewd, armless, legless creatures that spend their lives immersed in water. With their acute sonar and the sensitivity of their skin, they understand the world through hearing and touch to a degree that we cannot fully appreciate. Imagine being able to feel the motions of someone across the room. Moving effortlessly through the thin medium of air, we are almost oblivious to it. But for dolphins, water turbulence from storms, surf and their own motions is a palpable force they can readily exploit.

What, then, could be more natural—for dolphins, not humans—than to invent toys made of nothing but air and swirling water? With their innate sense of fluid dynamics and a little experience, blowing bubbles with complex shapes and movements is child's play. Except, of course, that human children can't play this way at all. It would be as though we could blow smoke rings, then use them as hula hoops.

Enjoy reading about this alien intelligence and marvel at how much we do—and don't—have in common with it.

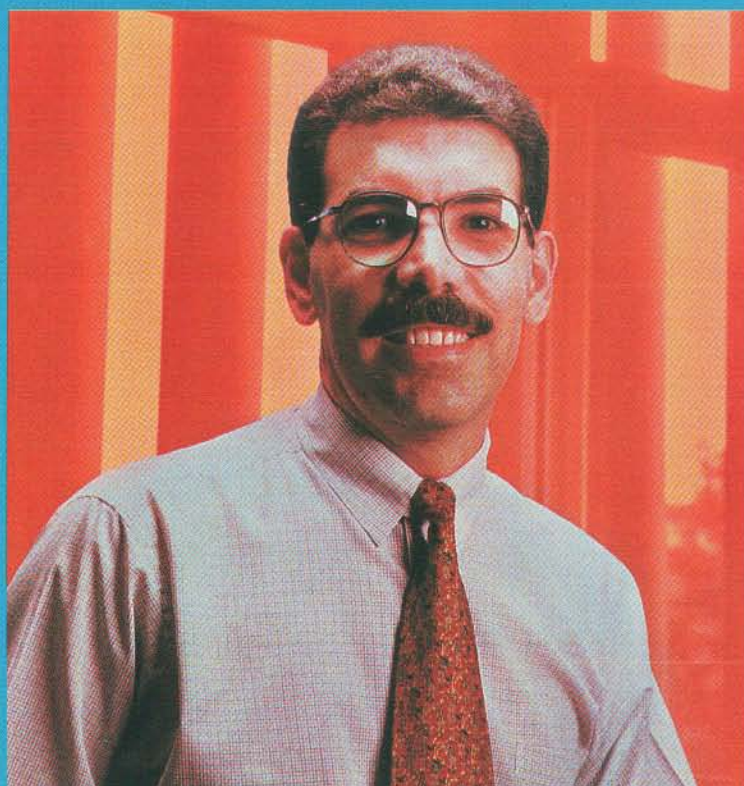
JOHN RENNIE, *Editor in Chief*
editors@sciam.com

SCIENTIFIC AMERICAN®
Established 1845

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue
New York, NY 10017-1111

PRINTED IN U.S.A.

Sometimes, simple things are the most important.



Let's make things better.

Walter Duran, Product Manager, Colour TV for Latin America.
Sao Paulo, Brazil.

Turning our limits into challenges, making breakthroughs.

At Philips, we do it everyday. That's how we can create better and simpler things.

I'm Walter Duran, product manager for colour TVs in Latin America, working in Sao Paulo.

I used to wonder, for example, how we could help people use their TV's potential, without making a bigger instruction manual.

Answer? A remote control device that puts instructions on the screen, so the TV's functions seem simple and instinctive. It's a method that's now adopted all over the world.

It wasn't easy, but it has made life easier.

PHILIPS

PHILIPS

LETTERS TO THE EDITORS

HELIUM FOR SALE

As a reader of *Scientific American* for two decades, I appreciated the article "No Light Matter," by Corey S. Powell [Science and the Citizen, March]. But the writer erred in claiming that President Bill Clinton and Congress will squander, rather than conserve, helium. The truth is, the government now squanders both helium and taxpayers' money. The Bureau of Mines's helium operation is $1.4 billion in debt, as it competes with private industry, which produces 90 percent of the world's helium.

In April the Helium Privatization Act of 1996, sponsored by Representative Barney Frank of Massachusetts and myself, passed the House. Under this legislation, the federal government will sell its helium operation and inventory—not for immediate consumption but to be maintained in the same underground dome in Texas where it is stockpiled. The helium will remain available for scientific and commercial use, just as it is today. One thing, however, will be different: the millions of dollars in annual losses will stop, and the $1.4-billion debt to taxpayers will be repaid.

CHRISTOPHER COX
Member, U.S. House
of Representatives
State of California

MEGA-DISCORD OVER NANOTECH

Congratulations on a fine Trends article by Gary Stix ["Waiting for Breakthroughs," April]. As much as I liked Richard Feynman's work, including his amusing 1959 lecture, I can't resist drawing parallels between the frequent appeals to the authority of Feynman by the nanotechnology crowd with similar claims by the cold-fusion mafia in the name of Nobel laureate Julian Schwinger. In his last years Schwinger became isolated from the mainstream scientific community, and shortly before his death, he wrote down some theoretical ideas about cold fusion. Thus, every cold-fusion propaganda piece drips with references to "Nobel laureate Julian Schwinger." Feynman gave his "nano" lecture at the height of his intellectual powers, but he did not intend to become a nano-Moses. Were he still with us, he would either vehemently reject the appeal to authority or, more likely, play along until he could turn it into a prank.

The motivation behind too much of the current promotion of nanotechnology can be summed up with a quote from the Foresight Institute Web site: "If you'd like a higher level of involvement, you may wish to join our Senior Associate program. By pledging an annual contribution of $250, $500, $1,000, or $5,000 for five years, you are brought into the circle of those most committed to making a difference in nanotechnology." I think that says it all.

JAMES F. HAW
Texas A&M University

I was dismayed to read an extended quotation from Feynman's essay "Cargo Cult Science" used as a critique of nanotechnology. I am sure he would have found such misuse of his idea quite objectionable. I should know because I talked with my father at length about the prospects of nanotechnology. As the article itself points out, Feynman saw no basis in physical laws that would preclude realization of the concepts of nanotechnology. To claim that nanotechnology is cargo cult science because its proponents analyze the capabilities of devices not yet constructed is as absurd as saying that astronautics was cargo cult science before *Sputnik*.

If my father were still alive, I think he would have been pleased to have his name associated with a large cash prize that seeks to accelerate the realization of one of his most exciting ideas. That is why I have participated in defining the conditions for winning the Feynman Grand Prize and have agreed to naming the prize in his memory.

CARL RICHARD FEYNMAN
Acton, Mass.

I am quite upset that a reference made in jest to the writer Stix was used out of context to ridicule nanotechnology and the conference we both attended. With a graduate degree in biomedical engineering as well as dentistry, I do not consider myself an "aesthete of science and technology."

EDWARD M. REIFMAN
Encino, Calif.

The article by Stix was a lengthy piece containing many errors and omissions. Your readers can find a critique of the piece with links to the broader literature at http://www.foresight.org/SciAmResponse.html, or they can send an electronic message to inform@foresight.org to request an e-mail version.

K. ERIC DREXLER
Institute for Molecular Manufacturing
Palo Alto, Calif.

The Editors reply:

Reifman, who was quoted as saying that Drexler is the messiah, maintains that his comment was made in jest. But he confirmed the sense of the quote when he was contacted for fact-checking prior to the article's publication. And with apologies to Drexler, we think that readers of the critique will find little in the way of specific cited errors.

BLINDED BY THE LIGHT

James Burke's column "Connections" is uniformly a pleasure to read, but I would like to call your attention to a small slip in his April piece ["What's in a Name?", Reviews and Commentary, April]. He correctly describes the Fraunhofer lines in the sun's spectrum—which are caused by atomic absorption—as dark lines. In the Kirchhoff-Bunsen flame, however, the lines are not dark but bright, as they result from atomic emission. I hope this mistake will fade unnoticed into oblivion, but for a spectroscopist, it is literally a glaring error.

GABOR B. LEVY
International Scientific Communications
Shelton, Conn.

Letters may be edited for length and clarity. Because of the considerable volume of mail received, we cannot answer all correspondence.

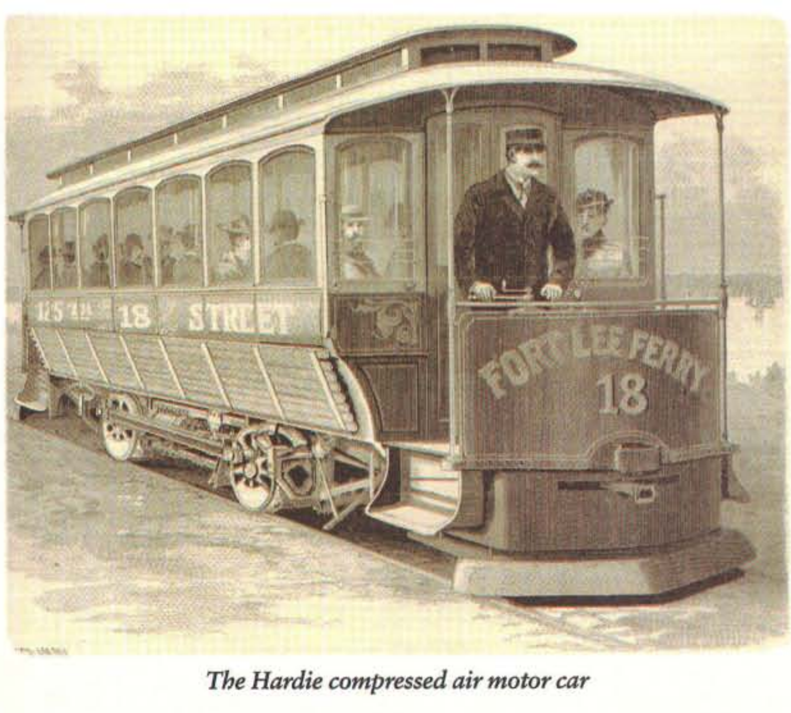
50, 100 AND 150 YEARS AGO

AUGUST 1946

According to one contention, magnesium will eventually replace iron as the world's basic constructional raw material. Hence, it might be feasible to call the next age of man the 'magnesium age.' The element appears to be the only 'basic' material of which the supply is inexhaustible: one cubic mile of sea water contains 9.2 billion pounds of metal in the form of magnesium chloride. It is the lightest of the structural metals, and magnesium's so-called 'fire hazard' is only a factor when handling fine powders or the molten metal. However, if magnesium is to become the prime raw material it is not likely to do so for centuries. Its competitors—iron and steel, aluminum and structural plastics—would have to reach a state of depleted supply and high prices."

AUGUST 1896

Interest in the compressed air motor has been shown by the Third Avenue Railroad Company, of New York, which has adopted the system invented by Mr. R. Hardie. In earlier systems, when the air was expanded from the storage flasks, the corresponding reduction of temperature was so great as to cause freezing and choking up of the exhaust passages. In the Hardie system, the cars, one of which is shown in the accompanying illustration, are similar in their general appearance to an ordinary street car. But underneath the seats are sixteen air reservoirs, rolled steel flasks 9 inches in diameter and 20 feet long, and a hot water tank, by means of which the air is heated before it enters the two cylinders of the motor, and the difficulty of freezing exhaust is overcome."



The Hardie compressed air motor car

"The Roentgen rays produced by the Crookes tube are now declared, by Nikola Tesla, to be material particles. Mr. Tesla states, 'The cathode stream is reduced to matter of some primary form heretofore not known.' "

"Dr. Fridjof Nansen, the Norwegian Arctic explorer, has attained the highest latitude yet in the quest to reach the pole, that of 86 degrees 14 minutes. Dr. Nansen says, 'At latitude 78 degrees 50 minutes north, we allowed our ship, the Fram, to be closed in by the ice. As anticipated, we drifted northwest during the autumn and winter. Lieut. Johansen and I left the Fram on March 14, 1895, to explore to the north and reach the highest latitude possible. We had twenty-eight dogs, two sledges and two kayaks for possible open water. However, by April 7 the ice had become so rough that I considered it unwise to continue.' They headed south and after a winter of living on bear and walrus meat in a stone house they had built, the two explorers were picked up by the steamer Windward on the coast of Franz Josef Land."

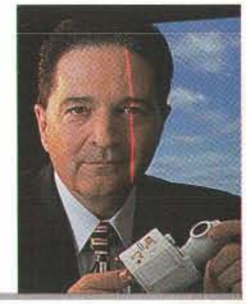
AUGUST 1846

By means of a magnificent and powerful telescope, procured by Lord Ross, of Ireland, the moon has been subjected to a more critical examination than ever before. It is stated that there were no vestiges of architectural remains to show that the moon is or ever was inhabited by a race of mortals similar to ourselves. The moon presented no appearance that it contained anything like the green-field and lovely verdure of this beautiful world of ours. There was no water visible—not a sea, or a river, or even the measure of a reservoir for supplying a factory—all seemed desolate."

"It is well known that there is a constant emission of hydrogen from the decomposition of various substances; and that this gas, being buoyant, has a tendency to rise to the surface of the atmosphere. According to one view, there is therefore no doubt that immense quantities of this inflammable substance abound in the upper regions, and that a spark of electric fire would envelope the world in flames. The only circumstance preventing such conflagration is that the region of excitable electricity is several miles below that of the inflammable air."

"Homæopathic soup: Take two starved pigeons, hang them up by a string in the kitchen window, so that the sun will cast a shadow of the pigeons in an iron pot on the fire, holding ten gallons of water. Boil the shadow over a slow fire for ten hours, and then give the patient one drop in a glass of water every ten days."

NEWS AND ANALYSIS



IN FOCUS

PAYING ATTENTION

The controversy over ADHD and the drug Ritalin is obscuring a real look at the disorder and its underpinnings

KATHERINE LAMBERT

GIRLS WITH ADHD
are increasingly joining the ranks of boys with the disorder—thereby contributing to the use of Ritalin in the U.S.

When Tom was born, he acted like a "crack baby," his mother, Ann, says. "He responded violently to even the slightest touch, and he never slept." Shortly after Tom turned two, the local day care center asked Ann to withdraw him. They deemed his behavior "just too aberrant," she remembers. Tom's doctors ran a battery of tests to screen for brain damage, but they found no physical explanation for his lack of self-control. In fact, his IQ was high—even though he performed poorly in school.

Eventually Tom was diagnosed with attention-deficit/hyperactivity disorder (ADHD)—a condition that typically manifests in young children as inattention or impulsivity and sometimes hyperactivity. These traits make it difficult for ADHD kids to sit still, concentrate and learn. The psychiatrist told Ann that in terms of severity, Tom was 15 on a scale of one to 10. As therapy, this doctor prescribed methylphenidate, a drug better known by its brand name, Ritalin.

Tom is now in fifth grade and lives with his father, Ned, and his problems have worsened. Ned has come to doubt that ADHD exists and took Tom off medication last fall. Many parents have in fact become suspicious of Ritalin after a recent surge in the number of children diagnosed with ADHD. By some estimates, as many as 5 to 6 percent of all school-age boys in the U.S. now take Ritalin for the condition. And production of the drug has shot up some 500 percent since 1990. Ninety percent of the current annual total, approximately 8.5 tons, is made by Ciba-Geigy and is used in the U.S.

Skeptics suggest that psychiatrists are too ready to diagnose a range of behavioral problems as ADHD and to dismiss them with a quick chemical fix. This past February the United Nations's International Narcotics Control Board reported that overdiagnosis of ADHD was very possibly taking place. In addition, the board declared that more teenagers were inhaling

the stimulant—which is related to cocaine but is far less potent—to get high. (Addiction is exceedingly rare.)

No one denies that abuse and misuse arise. Anecdotes abound about parents who seek an ADHD diagnosis for their child so that he or she can study more intently, take more time on tests and get better grades. Yet many of the pediatricians and psychiatrists treating ADHD kids believe the real explanation for the seeming increase in ADHD is far less complex: treatment is just now catching up to true prevalence. In the meantime, the media circus surrounding ADHD and Ritalin, they say, is hurting kids, like Tom, who need medication.

"The number of cases has more than doubled in the past five years, and so the chance that overdiagnosis is occurring needs to be considered," says James M. Swanson of the University of California at Irvine, "but even so, we are just now reaching the accepted range of the expected prevalence." Swanson and others cite several reasons why ADHD may have been previously underdiagnosed. First, physicians used to take children off medication when they reached adolescence for fear of long-term side effects. Now, though, most feel Ritalin is the safest psychotropic drug available and prescribe it even into adulthood. Also, ADHD was seldom recognized in girls before 1994, when the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*) noted a subtype of ADHD that appears without hyperactivity. ADHD girls are often not as antsy as affected boys, but they are restless mentally.

"Ritalin use is clearly more common now than ever before, and so people are saying that there is some implicit scandal afoot—that we are giving kids medication rather than dealing with their real problems," says Russell A. Barkley of the University of Massachusetts Medical Center. "But that's just blowing smoke." Edward Hallowell, a child psychiatrist at Harvard University who treats ADHD and has it himself, agrees: "This sort of criticism is just another example of what Peter Kramer, author of *Listening to Prozac*, calls psychopharmacological Calvinism." We live in a society that expects you to fix things yourself, he explains. Relying on any help, be it counseling or medication, is considered a weakness.

It will be difficult, though, to move from making moral diagnoses to medical ones because all the available tests for mental illness are so subjective. The criteria set forth for ADHD in the *DSM-IV* require that a child display a range of symptoms, such as distractibility and a short attention span, that are excessive for his or her mental age. Moreover, these symptoms must persist for at least six months and significantly impair the child's ability to function.

Nearly all children exhibit some of these symptoms some of the time. And ADHD falls along a spectrum, as do all psychological disorders. "Where we draw the line along that spectrum determines how many people have it," Barkley notes. Making diagnosis even more difficult is the fact that ADHD frequently appears with other disorders, including Tourette's syndrome, lead poisoning, fetal alcohol syndrome and retardation. In addition, many other conditions—such as depression, manic-depressive illness, substance abuse, anxiety and personality disorders—share similar symptoms.

Nevertheless, the biology behind ADHD is beginning to surface. "We cannot say which structure or which chemical is wrong," emphasizes Alan Zametkin of the National Institute of Mental Health (NIMH). "ADHD is like fever—any number of causes can be to blame." But he has found, for example, that a small subset of ADHD people have a different receptor for thyroid hormone and that 70 to 80 percent of all people with this very rare difference in their thyroid receptor have ADHD.

Other studies have found an association between ADHD and three genes encoding receptors for the neurotransmitter dopamine. Collaborating with molecular biologists and geneticists at Irvine and at the University of Toronto, Swanson examined the so-called novelty-seeking gene, which codes for the dopamine receptor DRD4. One series of base pairs repeats two, four or seven times. More repeats are associated with a blunted response to dopamine signals and less inhibited behavior. "We found that the seven-repeat variety of the gene is overrepresented among ADHD children," Swanson says.

Neurochemistry is not the whole story. Scientists have also discovered structural abnormalities. F. Xavier Castellanos of the NIMH used magnetic resonance imaging to measure the total brain volume and several different brain regions in 57 ADHD boys and 55 healthy control subjects. His team found that the anterior frontal part of the brain was on average more than 5 percent smaller on the right side in ADHD boys. The right caudate and the globus pallidus, too, were smaller. These structures form the main neural circuit by which the cortex inhibits behavior, and so damage there might well manifest itself as a lack of impulse control. Castellanos warns that this result offers but part of the puzzle: "It's only slightly better than phrenology. Now we're just measuring the bumps on the inside of the brain."

JEFF MERMELSTEIN

***PILL REGIMEN** for an entire family with ADHD includes daily doses of Ritalin.*

Another facet of ADHD malfunctioning comes from positron emission tomography (PET) studies. Julie B. Schweitzer of Emory University monitored brain activity in ADHD and unaffected men while they completed a task. Participants heard a series of numbers, one every 2.4 seconds, and were asked to add the last two digits they heard. Looking at the PET scans, Schweitzer saw two major differences between the groups. First, the ADHD individuals maintained high levels of blood flow, whereas the controls displayed deactivation in the temporal gyrus region—indicating some kind of learning.

The ADHD group also activated brain areas used for visual tasks. "I went back and asked the ADHD subjects if they used some strategy," Schweitzer says. "Instead of repeating the numbers to themselves, as some of the controls did, many ADHD patients had visualized them." She suggests that this visualization represents some kind of compensation for impaired cognition elsewhere. Zametkin, too, has used PET scans to study ADHD. He took images of parents of ADHD children and found that they exhibit less brain activity. He concludes, "These kids really are born to be wild." *—Kristin Leutwyler*

REMOTE SENSING

PUBLIC EYE

Spy satellite technology may assist government watchdogs

Last year President Bill Clinton signed an order declassifying hundreds of thousands of photographs taken by the first-generation of military spy satellites in a program that ended in 1972. Within another two years, commercial satellite companies plan to deliver pictures of better quality to anyone with a credit-card number and a Federal Express or an Internet account. They intend to sell snapshots from space that can show details as small as a meter—a close enough view to delineate boats, bridges or houses anywhere on the planet.

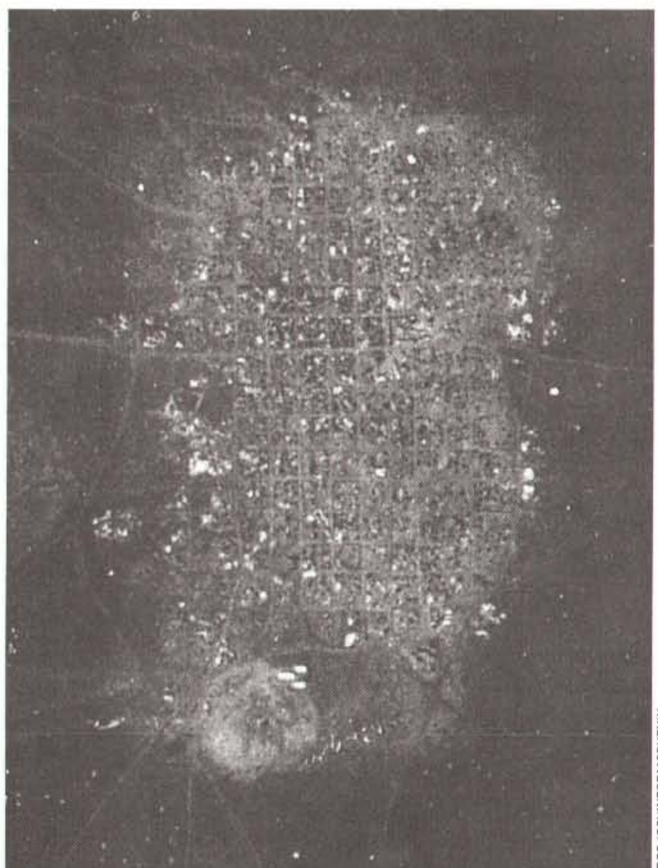
The companies have already publicized the imminent arrival of high-resolution satellite images as a boon for business. Real estate agents could furnish prospective buyers with a panoramic look at a neighborhood. Travel agents may provide vacationers with a dramatic overview of a chateau in the Alps.

But perhaps the most intriguing application for this erstwhile spy technology may be for public-interest groups and news organizations to keep an eye on government. "When one-meter black-and-white pictures hit the market, a well-endowed nongovernmental organization will be able to have pictures better than [those] the U.S. spy satellites took in 1972 at the time of the first strategic arms accord," comments remote-sensing and arms-control expert Peter D. Zimmerman. A case that immediately comes to mind is the stunning U.S. government satellite and spy plane images that showed a group of people herded onto a field near the town of Srebrenica in Bosnia and a newly dug mound of earth there that suggested the location of a mass grave. A public-interest group, unencumbered by internal policy debates, would likely move more quickly than a government in making similar pictures available.

Human Rights Watch communications director Susan Osnos remarks that satellite imagery could prove a valuable adjunct to on-site monitoring visits and testimonials from witnesses when investigating cases of rights abuses. "Last year when it became clear that more than 7,000 men were not going to reappear, we talked about the fact that there were all these surveillance satellites and that there must therefore be photographic evidence," Osnos says. "Had we been able to put our hands on the photos at that time it would have been a very powerful advocacy tool."

One organization, the Federation of American Scientists (FAS), has recently launched an initiative, called Public Eye, to promote the use of intelligence technologies, including one-meter imagery. "Information is power," says John E. Pike, an analyst with the FAS. "But before it was only available to a superpower. Now it will be available to any organization or individual for a few thousand bucks. This has the potential to expand the range of issues on which nongovernmental actors make news." (See the Public Eye page on the World Wide Web at http://www.fas.org/eye/)

The work of a Norwegian graduate student, Einar Bjørgo, presages how remote sensing may help international relief efforts. Bjørgo, a student at the Nansen Environmental and Remote Sensing Center at the University of Bergen, has used 1992 images with two-meter resolution from a Russian spy satellite to show how the size of refugee camps in the Sudan can be estimated. (The images can be accessed on the Web at http://www.nrsc.no:8001/~einar/UN/refmon.html) Bjørgo obtained the pictures from a Russian company that has marketed slightly out-of-date satellite imagery for several years.



NERSC/SOVINFORMSPUTNIK

SATELLITE IMAGE ANALYSIS of a Sudanese refugee camp enabled a Norwegian institute to develop a method of estimating a site's area and population.

The news media will also benefit from improved views from on high. For the past decade, some journalists have offered more incisive coverage with satellite pictures. ABC News has merged less distinct satellite images with digital map information to create computerized landscape representations for stories on the Persian Gulf War or North Korean nuclear facilities. But Mark Brender, a producer at ABC News, still laments not having access in 1990 to high-resolution images, which would have shown Iraqi tank columns moving into Kuwait. Lower-quality pictures, procured by ABC from the Russians, were not enough to elicit the necessary detail.

These same images, with roughly five-meter resolution, did provide enough information for remote-sensing expert Zimmerman to ascertain that overall Iraqi troop buildups had been overstated by the Bush administration, a fact subsequently acknowledged by government offi-

cials. Zimmerman was working under contract to the *St. Petersburg Times,* which published a story on his findings. One-meter imagery would have made his job much easier. "I would have been able to make conclusions with extremely high confidence," he declares. "I would have been able to see individual vehicles on the road."

The growing interest in satellite news gathering has gained enough momentum for American University professor Christopher Simpson to set up the Project on Satellite Imagery and the News Media at the university's School of Communication. The group has put together guides for journalists that contain legal background relating to satellite imagery usage and public sources of satellite data available on the Internet. (The guide to remote-sensing data can be found on the Web at http://grukul.ucc.american.edu/earthnews)

Whatever the uses, the future of one-meter imaging will depend on a successful launch by at least one of three companies—Space Imaging, Orbital Sciences and EarthWatch, all of which plan during the next two years to put up satellites that will circle the earth at an altitude of a few hundred miles. The fate of the high-resolution commercial market will also rely on a measure of government leniency.

The National Oceanic and Atmospheric Administration, which licenses commercial satellites, held hearings in mid-June on updating a regulation that gives the government broad latitude in imposing "shutter control"—that is, the right to restrict satellite data deemed to compromise national security or foreign policy. Media representatives want satellite pictures to be guaranteed First Amendment protections that will make it difficult to bar access to the images. Besides domestic limits, satellite companies may have to contend with blackouts imposed overseas. Citing national security, Israel has reportedly asked the U.S. government to restrict the resolution of detail in commercial satellite pictures of its territory to no less than three meters. Only time will tell whether governments get a bad case of cold war feet. —*Gary Stix*

FLUID DYNAMICS

THE WALL FALLS

A half-century-old equation for fluid dynamics is in doubt

Alexandre Chorin and Grigory Barenblatt had been studying turbulence from different perspectives for more than 30 years when they met this past February at the University of California at Berkeley. Chorin works in computational fluid dynamics, calculating the theoretical properties of idealized turbulent flow. Barenblatt is a mathematician who studies the "scaling laws" that engineers employ to extrapolate results from wind-tunnel tests and other small-scale experiments to the real world.

But the two saw ground for collaboration: theoretical studies of turbulence have been limited for some years by the mathematical formulations of fluid flow. Even after extensive refinement of experimental apparatus, discrepancies remained between predicted results and actual measurements. The only way to go forward was to go back and reexamine the foundations of the field, Barenblatt recalls.

The foundation they revisited was the Law of the Wall, an equation formulated in the 1930s by Theodor von Kármán to describe the forces that turbulent flows exert on solid objects. In doing so, Barenblatt claims, von Kármán made a simplifying assumption that seemed so obvious no one questioned it for nearly 50 years: while investigating the viscosity, or resistance to flow, caused by turbulent eddies, von Kármán and others ignored the minuscule viscosity added by the random thermal motion of individual molecules.

This tiny molecular viscosity sometimes has disproportionate effects. When Chorin and Barenblatt rederived the law to take the jostling molecules into account, they found that under some conditions—particularly, at higher speeds and pressures—the force exerted by a turbulent flow was significantly higher than that predicted by the old equation. The new version's predictions for the transfer of heat in a turbulent flow differ even further from earlier ones. In a way, molecular viscosity behaves like the notorious butterfly wing of chaos theory, whose delicate flapping could trigger a chain of events leading to monsoons half a world away.

Reaction to the new formula has been mixed. Older fluid dynamists have spent

IN BRIEF

The Other White Fish

Sea lamprey. These slimy, eellike parasites normally suck the life out of trout and salmon fisheries in the Great Lakes. The Great Lakes Fishery Commission traps 50,000 to 100,000 lamprey every year—sterilizing and releasing the males and sending the females to the landfill. But researchers from the Sea Grant Program at the University of Minnesota at Duluth have a new plan: sell them to the Portuguese! There lampreys are considered a tasty meal. Sea Grant will send a sampler of 80 females overseas this summer.

ZIG LESZCZYNSKI *Animals Animals*

Making Memories

As people age, an enzyme called prolyl endopeptidase (PEP) increasingly degrades the neuropeptides involved in learning and memory. In Alzheimer's disease and senile dementia, the process is accelerated, causing memory loss and a shortened attention span. But now researchers in Suresnes, France, have found compounds that prevent PEP from breaking neuropeptides apart. In tests, these chemicals almost completely restored memories in amnesiac rats.

Destruction of Smallpox Postponed...

Until the summer of 1999 at least, now say officials from the World Health Organization. The killer bacterium was eradicated in 1977, but samples of it have remained under guard in the U.S. and in Russia. The new deadline for destroying those remaining vials is in fact the third to be set. Two earlier dates passed while scientists debated the value of thoroughly studying the virus's genetics before eliminating it.

Pesticides on the Rise

A draft report from the Environmental Protection Agency issued this past May states that the use of active pesticide ingredients rose from 1.23 billion pounds in 1994 to 1.25 billion pounds in 1995. Many environmental groups fear the numbers are somewhat misleading because the EPA did not take into account inert ingredients, wood preservatives or disinfectants, which can also be toxic.

Continued on page 12

In Brief, *continued from page 11*

Case Closed

After 84 years, the Piltdown hoax may be solved. In 1912 Arthur Smith Woodward—keeper of paleontology at London's Natural History Museum—hailed bones from Piltdown, England, as the Missing Link. But some 50 years later it became clear that a criminal—and not evolution—had joined the human skull and orangutan jaw. Recently two scientists analyzed similarly stained specimens in an old trunk bearing the initials M.A.C.H. and, at last, fingered the perpetrator: Martin A. C. Hinton, a curator of zoology, who had warred with Woodward over wages.

UPI/CORBIS-BETTMANN

Cooperative Crustacea

A new study shows that *Synalpheus regalis*—snapping shrimp that dwell within sponges on Caribbean coral reefs—are eusocial. Colonies contain a single reproductive female and workers that help to defend her. Many eusocial creatures, such as bees and ants, are a haplodiploid species—that is, males develop from unfertilized eggs and females from fertilized ones. But *S. regalis* males and females both come from fertilized, or diploid, eggs—as do naked mole rats, another eusocial creature. The discovery marks the first case of eusociality in crustaceans.

Budget Woes

The American Association for the Advancement of Science (AAAS) has recently analyzed the budget plans for fiscal year 1997 put forth by the president and the House of Representatives. Both proposals, the AAAS says, mean a reduction of nearly 25 percent by the year 2002 in nondefense research and development—a dramatic cut.

Letting Loose

The National Aeronautics and Space Administration and the Italian Space Agency have at last released a 358-page report explaining why, on the space shuttle *Columbia*'s most recent mission, the Tethered Satellite broke free. Something caused a break in the insulation surrounding the tether's conductor. Current from this copper wire then jumped to a nearby electrical grounding site. The current burned through much of the tether until, finally, it snapped.

Continued on page 14

their careers adhering to the old Law of the Wall, Barenblatt says, and some of them are unwilling to see it pass without rigorous examination of its replacement. If it is confirmed, the revised version could have significant implications for systems as disparate as industrial heat exchangers and global climate models. Boilers, air conditioners and other devices might have to be redesigned with new proportions to improve their efficiency and lengthen their working lives.

Luckily, one application of fluid dynamics that will be less affected is aircraft design—aeronautical engineers have never used the Law of the Wall directly, Barenblatt says. (Instead they have relied on extensive experimental data backed by tried and true scaling rules.) What the new law may do, he predicts, is to make turbulence easier for many engineers to understand—creating a smoother flow from theory to practice. —*Paul Wallich*

ECOLOGY

MUSSEL MAYHEM, CONTINUED

Apparent benefits of the zebra mussel plague are anything but

When the first zebra mussels were spotted in Lake Erie in 1988, the Cassandras had a field day. Within weeks, there were predictions that the incredibly hardy, prolific creatures would bring on ecological and financial disaster as they wreaked havoc with the lake's food chain and clogged the water-intake systems of electric power stations, boat motors and drinking-water facilities.

In fact, in many ways the disaster did not live up to expectations. Chlorine and other chemicals have kept the pesky mollusks away from intakes at far lower costs than were feared. Moreover, there have even been some apparent benefits. The fat little critters are prodigious filterers, absorbing surprising amounts of a variety of pollutants from the water and storing them in their lipids. They have also consumed so much algae, their main food, that large parts of the lake have become visibly clearer.

Previously scarce aquatic plants, which were fighting a losing battle to the pollution-nourished algae, are thriving once again. Eel grass—an indigenous plant that Ohio was on the verge of declaring endangered in 1988—has rebounded so thoroughly that huge, tangled underwater forests of it now gently sway over the lake bed. The vegetation snags on the propellers of pleasure boats near Put-in-Bay, a tourist town on South Bass Island in the lake's western basin. Can it be that the zebra mussel

GARY MESLAROS *Bruce Coleman, Inc.*

ZEBRA MUSSEL BEACH
on the Great Lakes has come to symbolize the battle lost against this ecosystem-altering invader.

Alas, no, say local zoologists studying the infestation. Yes, zebra mussels have absorbed so much pollution that experts now estimate that 50 percent of the contaminants that had been in Lake St. Clair, a 1,200-square-kilometer body of water between Lake Erie and Lake Huron, are now in zebra mussel tissue. And, yes, in great sections of Lake Erie, the water is 600 percent clearer than it was. But if these facts seem like bright spots in a dark cloud, they are more lightning than silver lining.

Zebra mussels are living filters, says Susan W. Fisher, a professor in the entomology department at Ohio State University. Each adult mussel sucks in as much as a liter and a half of water a day, retaining algae and other nutrients and, incidentally, PCBs, dioxins, polynuclear aromatic hydrocarbons and whatever other contaminants happen to be in the water and sediments where the mussels feed. (The current population filters the entire western basin of Lake Erie every five to seven days.) Some of the toxins wind up in the animals' fat tissue—the mussels are a rather ample 15 percent fat by weight, Fisher notes.

Billions of mussels with contaminated fat may not seem like a big problem, but ecologists are concerned. For example, the mussels' corpses and feces, which are also contaminated, are important links in a food chain extending to anyone who eats fish from the Great Lakes. "The point is, it may have been safer to have the contaminants in the sediments," says Jeffrey M. Reutter, who is director of Ohio State's Stone Laboratory, the oldest freshwater biological field station in the U.S.

Fisher is now trying to determine whether the contaminants are reaching dinner tables in concentrations high enough to be troubling. "It may take a couple of years to know if there's a wholesale rearrangement of contaminant concentrations going on in the lake," she says.

The mussels' reduction of artificially high levels of some kinds of algae in Lake Erie may also have had devastating repercussions. Algae is the base of the food chain for all the lake's creatures, so its rapid loss on such an enormous scale may have caused fundamental changes, Reutter and Fisher suggest. For example, not only zebra mussels but also zooplankton subsist on algae. And the lake's food fish—walleyes, bass, trout and yellow perch—eat zooplankton at critical points in their lives but very rarely consume zebra mussels.

Before the infestation, Lake Erie, which is by far the most heavily fished of the Great Lakes, supported fisheries with an economic value of $600 million a year. During the early 1990s, with food fish apparently much scarcer, the value was down to $200 million a year. Reutter comments, however, that scientists have not yet conclusively linked zebra mussels to declining zooplankton populations or fisheries or to a few other baffling phenomena, such as the appearance of gigantic blooms of certain toxic algae species not eaten by zebra mussels. Scientists are also trying to determine if these and other changes have endangered the lake's entire ecosystem.

Whether the invertebrate intruder is ultimately blamed or not, Fisher has already reached one conclusion on her own. "Every little benefit you get out of them," she warns, "is not worth the problems." *—Glenn Zorpette*

ANTI GRAVITY

Patient, Smell Thyself

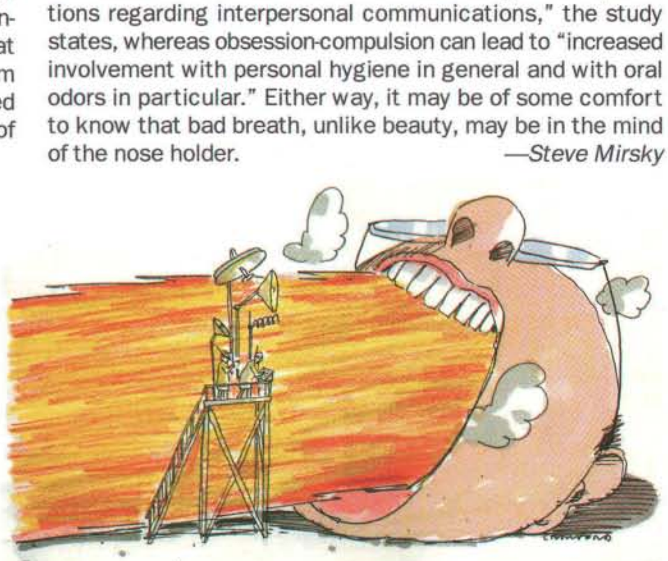
Worried that you have bad breath? Unless vultures are actually circling your mouth, there's good news. The problem may not be your breath at all, but your personality.

Researchers at Tel Aviv University decided to study just how bad the breath really was of 38 people whose concerns about their oral malodor drove them to seek medical attention. The researchers published the results in a recent issue of *Psychosomatic Medicine*. Sixteen patients had themselves come to the conclusion that they had a problem. Another 12 were driven to this conclusion—they claimed that others had complained. The last 10 were getting input from both sides, having decided for themselves that they reeked but having also found the telltale gift-wrapped bottle of mouthwash in their desk drawer.

As part of the study, the 38 subjects rated their own breath on a scale of zero to 10, where zero was presumably something like minty roses and 10 must have been whatever Linda Blair ejected onto the priest in *The Exorcist* that made him lose his faith. They also mouth-breathed from a distance of 10 centimeters right into the face of an "odor judge," who similarly rated the scent from zero to 10. To put the whole thing in perspective, the odor judge produced a baseline bad-breath value by assigning a rank rank for a control sample: dung-based fertilizer. The study subjects likewise rated the fertilizer, to prove that they did not suffer from anosmia—loss of sense of smell.

When the dust settled and the bodies were carted off, the ratings got analyzed, leading to some fascinating insights. Both the odor judge and the subjects rated the fertilizer at about nine on the stink scale. But whereas the odor judge rated the subjects on average to be far closer to mint than to manure, at 2.7, the study group assigned itself an average score suitable to grow a decent corn crop with—6.7.

Because the patients completed a psychological profile, the researchers were able to note higher than normal values for interpersonal sensitivity and obsession-compulsion. Increased interpersonal sensitivity may cause some to blame breath for their "self-consciousness and negative expectations regarding interpersonal communications," the study states, whereas obsession-compulsion can lead to "increased involvement with personal hygiene in general and with oral odors in particular." Either way, it may be of some comfort to know that bad breath, unlike beauty, may be in the mind of the nose holder. *—Steve Mirsky*



MICHAEL CRAWFORD

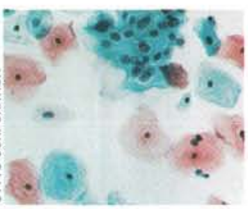
In Brief, *continued from page 12*

Name That Bug

To identify annoying garden pests, northern Californians can now call 1-900-225-BUGS between 10 A.M. and 4 P.M. Set up and staffed by members of the entomology department at the University of California at Davis, the hot line costs $2 for the first minute and $1 for each additional minute. This program is modeled after a successful one in Minnesota. And remember, you must be infested to call.

Catching Cervical Cancer

The Food and Drug Administration recently approved a better method for cervical cancer screening. The disease, which kills some 4,900 women annually, is highly treatable when caught early. To examine cervical cells, doctors have traditionally smeared a tissue sample—containing blood cells and mucus—against a glass slide (a Pap smear). In the new technique, cervical cells are filtered from the tissue sample first and then applied in a thin layer to the slide, making detection far easier.



CYTYC CORPORATION

FOLLOW-UP

Ozone Depletion Decreasing

Chemists at the National Oceanic and Atmospheric Administration say the average concentration of ozone-depleting chemicals in the lower atmosphere is falling off fast. Based on their calculations, the amount had decreased some 1 percent by the middle of last year. The decline suggests that the Montreal Protocol—a treaty banning the production of CFCs and other halogenated compounds—is having a real effect. (See September 1995, page 16.)

Canned Software

According to a letter written in June by Congressman Floyd D. Spence, chairman of the House National Security Committee, the U.S. Army will terminate its Sustaining Base Information Systems program at the end of fiscal year 1997. The program was to have replaced some 3,700 computer systems by 2002. To date, the army has spent more than $150 million yet has received only a handful of replacement systems. (See April 1996, page 26.)

—Kristin Leutwyler

ASTRONOMY

THE BLUSTERY VOID

Space weather forecasting comes of age

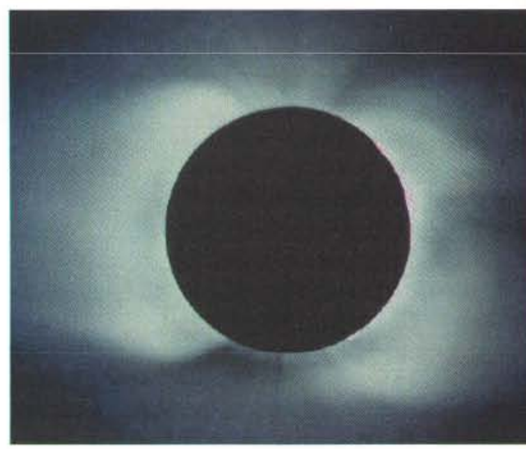
On March 26 of this year, the *Anik E1* telecommunications satellite lost power in one of its solar panel arrays, temporarily interrupting voice, video and data service for its owner, Telesat Canada. In the past, such a mishap might have been vaguely attributed to component failure. But this time, Daniel Baker of the University of Colorado identified a more specific culprit: a bout of inclement space weather.

Outbursts of magnetic flux and charged particles from the sun episodically roil interplanetary space and agitate the earth's magnetic field. These disturbances have long been known to induce surges in power grids and to interfere with long-distance navigation and communications signals. Improved understanding of space weather is revealing the true magnitude of the problem, as experienced by *Anik E1*. Ernest Hildner, director of the Space Environment Center (SEC) in Boulder, Colo., warns that the situation is only going to get worse. With their miniaturized circuits and reduced overall size, modern satellites are increasingly vulnerable, even as their total number continues to grow. Meanwhile the sun is likely to turn ever more restless as it progresses through its current 11-year activity cycle.

To help ameliorate the potential losses, the SEC collects data on the space environment around the clock. It relies in part on a new generation of scientific probes—most notably *POLAR, WIND* and *SOHO,* spacecraft residing upstream of the earth in the solar wind—that monitor the behavior of the sun and relay information about conditions in interplanetary space. The SEC can make predictions about general space weather up to three days in advance. Anyone doubting how far the study of space weather has come need only visit the SEC site on the World Wide Web (http://www.sel.noaa.gov/), which contains detailed, constantly updated records and forecasts.

The SEC is now working in collaboration with the National Weather Service's "weather wire" to warn via radio of severe space weather, much as the weather service would send out an alert for a hurricane or tornado. In addition, the National Science Foundation recently organized a National Space Weather Program to coordinate and disseminate research from several agencies, primarily the U.S. Air Force and the SEC.

A key goal of that program is to make space weather information available in a format that is useful to the companies and individuals who could most benefit from it. "Right now there is a problem in reporting to the public," notes Captain Amanda Preble, chief of space weather programs at the air force's Directorate of Weather. "This is not like a tornado that you can show people." That situation is beginning to change; the rich data streams coming from the spacecraft will soon make it possible to construct interactive, three-dimensional models of space weather.



ASTRONOMICAL IMAGE LIBRARY

SOLAR CORONA, the ghostly glow around the eclipsed sun, shows where space weather begins.

As a result, engineers facing turbulent space weather can think more carefully about their options. Utility companies are starting to monitor geomagnetic activity and may set aside additional reserve capacity during solar storms; cellular telephone companies could warn customers about potential transmission failures; in extreme cases, operators might place satellites in "sleep" mode or prepare to retransmit software commands that could be lost in a storm of charged particles. "People used to laugh, to consider our work very 'Star Trek,'" Preble recalls, "but it is already proving to be useful."

—Corey S. Powell

REPRODUCTIVE BIOLOGY

MOUSE TO FATHER RAT?

Renewable reproductive cells could transform fatherhood

"Mary had a little lamb—and the doctors were surprised," runs a warped version of the nursery rhyme analyzed by the linguist Steven Pinker in his book *The Language Instinct.* Today the doctors would still be very surprised. But the prospect of a mouse siring a rat, at least, has suddenly become an imminent reality, thanks to a technique developed at the University of Pennsylvania. The process—transplantation of cells that produce sperm—could also allow mammals, including humans, to father multiple offspring years after their death.

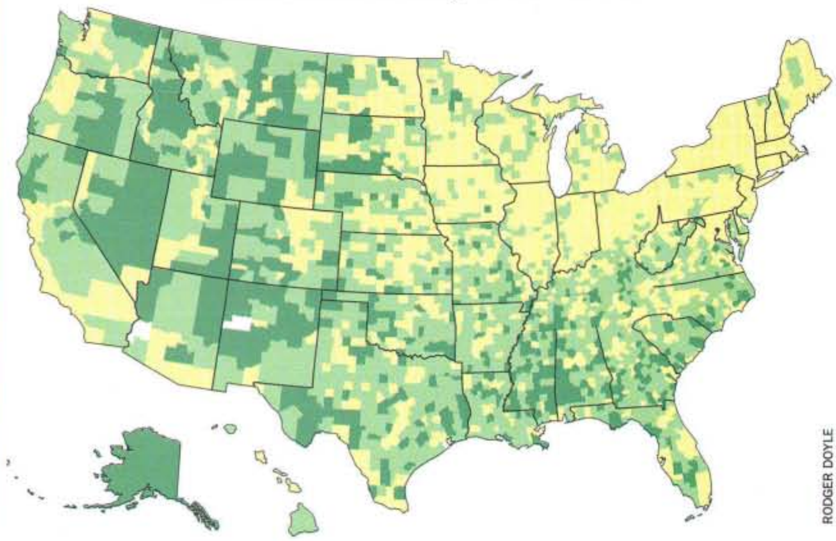
The researchers describe in the May 30 issue of *Nature* how they transplanted sperm-producing cells called spermatogonia from rat testes into mice testes, where the cells made seemingly normal rat sperm that may be capable of fertilizing rat eggs. A companion paper in the June issue of *Nature Medicine* reports that the spermatogonia can be deep-frozen in liquid nitrogen for long periods—156 days, so far—before they are successfully implanted. The transplanted cells were genetically marked to prove that they were indeed the progenitors of sperm in the recipient testes. The investigators suppressed the immune systems of the mice so they would not reject the foreign rat tissue.

These findings, made by a group headed by Ralph L. Brinster, are being hailed as a breakthrough that could have far-reaching applications in medicine, livestock breeding and the preservation of endangered species. They are all the more remarkable because they are, by the standards of modern biological research, fairly simple. The microinjection of the spermatogonia into the mice was done by Mary R. Avarbock, a postgraduate student in Brinster's lab. Recognizing the implications, the university has moved quickly to file patent applications.

Veterinary and medical researchers learned decades ago how to freeze and store sperm for later use. But thawed sperm often loses much of its fertilizing capacity. The spermatogonia used in Brinster's experiments include stem cells

BY THE NUMBERS

Lethal Accidents, 1979–1992



SOURCE: Centers for Disease Control and Prevention; county data for Alaska not available

AGE-ADJUSTED RATE PER 100,000 PEOPLE
UNDER 45 | 45 TO 64.9 | 65 OR MORE
ACCIDENTS

Accidents do not occur at random. People 85 years of age and older are 22 times more likely to die accidentally than are children five to nine years old. The risk for Native Americans is four times that for Asian-Americans and twice that for white Americans or African-Americans. Males suffer accidents at more than twice the rate of females, in part because they are more prone to risky behavior. Alaskans are more than three times as likely as Rhode Islanders to die in an accident. Texans are 21 times more likely than New Jerseyites to die in a natural disaster. Among the 100 most populous counties, Kern County, California (Bakersfield), has an accident fatality rate three times greater than Summit County, Ohio (Akron).

Accidents happen more often to poor people. Those living in poverty receive inferior medical care, are more apt to reside in houses with faulty heating and electrical systems, drive older cars with fewer safety features, and are less likely to use safety belts. People in rural areas have more accidents than city or suburban dwellers because farming is much riskier than working in a factory or office and because emergency medical services are less readily available. These two factors—low income and rural residence—may explain why the South has a higher accident rate than the North. The high rate in the Mountain States is the result, in part, of the rural nature of the region. Alcohol is an important contributor to many accidents, including not only car crashes but also falls, fires and drowning.

Almost 90,000 Americans die in accidents every year. In 1992, 47 percent died in motor vehicle collisions, 15 percent fell to their death, 8 percent inadvertently poisoned themselves (typically with legal drugs), 5 percent perished in fires (mostly house fires), 4 percent suffocated or choked to death, another 4 percent drowned, and 3 percent died because of a medical mishap (usually during surgery). Occupational fatalities—primarily involving vehicle crashes, falls and dangerous machinery—accounted for 5 percent or more of all accidental deaths. Sport and recreational accidents, which occur mostly during swimming and boating, accounted for 7 percent.

In 1995 the accident rate was less than half that of 1930 despite the huge growth in the number of old people, the most accident-prone group. The death rate from motor vehicle collisions has declined by about 40 percent since 1930, whereas the rate for other types of accident fell by two thirds. Most of the decline results not from changes in people's behavior but from better safety procedures and devices, such as improved burn treatment, seat belts, smoke detectors, nonflammable sleepwear for children, window guards in apartment houses, and superior highway design. —*Rodger Doyle*

that may represent an inexhaustible supply. Moreover, they appear to be fairly robust, which is one reason why researchers are rhapsodizing about the possibilities raised by Brinster's research. Men facing chemotherapy that causes sterility may be able to bank some spermatogonia, then have them reimplanted later. Spermatogonia from an endangered species might be used to generate sperm in the testes of a more common species. The resulting sperm could be used to create embryos via in vitro fertilization, and the embryos might then be implanted into the uterus of a closely related foster species.


JOY SPURR Bruce Coleman, Inc.

MOUSE
could host the sperm-producing cells of a rat.

Another tantalizing possibility is that spermatogonial stem cells from many species might be susceptible to gene targeting, a technique in which biologists design molecules to "knock out" a specific gene in a cell that is incorporated into a developing animal. Some of the animal's offspring lack the targeted gene entirely. Gene targeting can be done now in mice, but "it is a problem with other species," says David E. Clouthier, a member of the Pennsylvania team who is now at the University of Texas Southwestern Medical Center. The difficulty is that gene targeting makes use of embryonic stem cells, which have proved impossible to isolate in livestock and in humans.

If spermatogonia from most mammals can be propagated and used for gene knockouts, livestock breeding may never be the same again. And if the technique works in humans, artificial genetic manipulation of the germ line—long the province of science fiction—will suddenly become a perfectly feasible prospect. Addressing the ethical implications of the new developments in reproductive biology "is an ongoing process," Clouthier says. Not many years ago ethicists were declaring that manipulation of the germ line was unthinkable, but that, in any case, it wasn't feasible in the foreseeable future. That future now seems a lot more foreseeable.

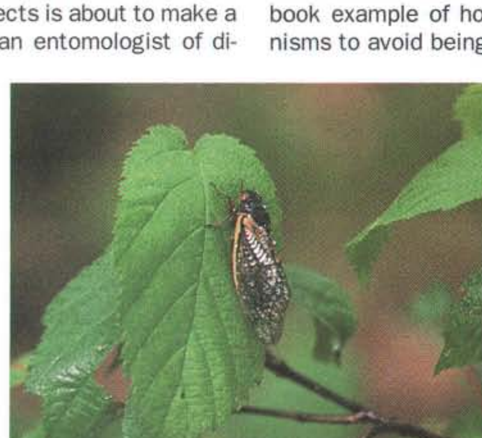
—Tim Beardsley in Washington, D.C.

FIELD NOTES

Insects of Generation X

A dozen reporters gather in New Haven, Conn., at Yale University's Peabody Museum of Natural History to hear Charles L. Remington speak in animated tones about the momentous event that is under way. In a nearby park, a group of astonishingly long-lived insects is about to make a rare appearance. And Remington, an entomologist of diverse interests, is doing all he can to enlist the help of journalists to ensure that people far and wide take notice.

The insect in the spotlight is the cicada. But the entomological star is not the ordinary, late-summer variety (the so-called dog-day cicada that noisily visits suburban backyards every year). The focus is on *Magicicada septendecim,* a species that is unique to the eastern U.S. *M. septendecim* lives most of its 17-year life underground, tapping fluids from tree roots for sustenance. With an uncanny sense of timing, these insects dig their way to the surface in the late spring of their 17th year, shed their final nymphal skin and populate a patch of forest. During their few weeks of adult life, they mate, deposit eggs in twigs and provide a feast for the birds that happen to notice that something special is occurring.


DAVID SCHNEIDER

The emergence of such a brood is indeed quite something for local predatory birds—and for scientists. Returning to the cicada colony after the formal media tour has ended, I encounter Stephen A. Marshall, an entomologist from the University of Guelph in Ontario, who has driven for 20 hours to visit the site. Marshall's research involves fly systematics, but he nonetheless desired to see this extraordinary phenomenon firsthand. He likens the event unfolding around us to other classic wonders of the insect world, such as the bioluminescent glowworm of New Zealand or the sacred scarab beetle of Egypt.

"It's a big, pretty animal that appears once every 17 years," notes Marshall, in an attempt to convey some insect aesthetics. He reminds me that this species is a textbook example of how insects adopt complicated mechanisms to avoid being preyed on. Cicadas are a particularly enticing food for birds, which tend to gorge on them: "We've had several reports of birds vomiting cicadas," Remington explains, "like a little kid with Grandma's cookies." Periodical cicadas, which live for either 17 or 13 years (depending on location), appear to have evolved their underground existence in an effort to escape being devoured into extinction. The strategy relies on outliving one's foes. Should bird populations expand as a result of the availability of these meaty insects, the predator's numbers will then fall over the next decade or more while the cicadas remain safely underground.

Watching the cicadas crawl over scattered shrubs, I find it difficult to acknowledge fully the longevity of these insects. (This species is perhaps the longest-lived insect in the world.) Indeed, I must constantly remind myself that the creatures drawing our attention are as old as many of Yale's incoming students. Marshall, positioning himself to take a close-up photograph, coaxes his young son not to disturb a nymphal cicada that has just emerged from the ground. Yes, Alexander, I murmur to myself, do show some respect for your elders.

—David Schneider

PALEONTOLOGY

A New King and His Tiny Minion

Poor *Tyrannosaurus rex* has been dwarfed, again. Fossils unearthed in the Kem Kem region of Morocco point to the existence of a dinosaur whose head was five feet, four inches long (1.6 meters), just slightly larger than that of *T. rex*. The discovery of *Carcharodontosaurus*, or "shark-toothed reptile," by Paul C. Sereno of the University of Chicago and his colleagues comes right after the finding last year of *Giganotosaurus* in Argentina. The South American giant and its new African counterpart—along with Sereno's other Moroccan find, a smaller species called *Deltadromeus*, or "delta runner"—are also helping scientists understand exactly when the continents split apart.

Paleogeographers believe that by the end of the Jurassic, some 150 million years ago, the ancient supercontinent Pangaea split into a section called Laurasia, which moved north, and Gondwana, which remained in the south. This idea is supported by fossils showing an evolutionary schism: species unique to each landmass sprung up at about the same time. But until now, this evidence had been restricted to Asia, Europe and South America.


PAUL C. SERENO

The Moroccan bones—the first major dinosaur fossils to be dug up in Africa—provide data suggesting that Pangaea's initial subdivision was not complete. *Carcharodontosaurus* and *Deltadromeus* both appear to have lived during the Upper Cretaceous, approximately 100 million years ago. Because this date is 50 million years after the purported Laurasia-Gondwana divide, scientists expected the African dinosaurs to be more closely related to those from the southern continents, such as South America. That, however, is not what they found.

Both African species are very similar to dinosaurs that roamed what is now North America some 100 million years ago. This discovery, along with the age of the new fossils, suggests that land bridges and shallow seas between Laurasia and Gondwana allowed dinosaur species to intermix throughout the Upper Cretaceous—some 90 million years ago and 60 million years later than was thought. *—Gunjan Sinha*

CYBER VIEW

The Notworking Computer

The network computer is almost a good idea—but not quite. Now widely touted as the next hot thing from the computer industry, the network computer is in fact shaping up to provide a classic example of how engineers get things wrong. The case for the network computer, or NC, as it is cozily called by big-name boosters such as Oracle, IBM, Apple and Sun Microsystems, is based on perceptive technological analysis, which veers unerringly to the wrong conclusion. Most visions of networked computing currently base their advantages on the convenience of builders and maintainers of computers. Those that actually sell will have to appeal instead to the convenience of users.

The touted advantages of the NC stem from the fact that it is half a computer. The user buys only screen, microprocessor and keyboard, not the disk on which programs and data are stored long-term. So the purchase price is relatively cheap—about $500. A (preferably high-speed) network link connects the user to the centralized disk storage on which programs and personal data are kept, as well as to the Internet. Because programs are stored centrally, the design minimizes the costs and complexity of managing machines—upgrading software, backing up data and so on—which can cost up to $3,000 a year at some large companies.

Technical specifications for NCs were released in May, and machines are expected on the market by Christmas 1996. The nice, hopeful aspect about the technology is that it is based almost entirely on open networking standards—those of the Internet and the World Wide Web. Machines meeting the specifications must also be able to cope with common audio, sound and video formats.

Most important, the machines must understand Sun's Java programming language, which enables small programs, called applets, to be sent over the network. In effect, Java makes it possible to have software on demand—and not just on demand but also delivered just in time. If your word processor, for example, does not understand the format of the document sent to you by some far-flung colleague, it could quickly and automatically download a Java applet from the Internet to do the job.

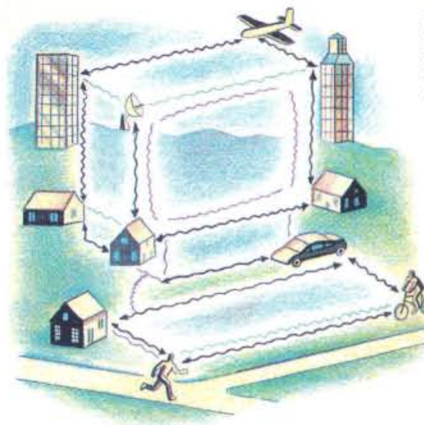
The flash of insight that inspired the NC is that this combination of technology enables the computer to be, in effect, deconstructed. Because different machines can work together over the Internet, no single machine need do everything a user requires. Instead it can call on other machines when it faces a task it can't cope with. But, having had this insight, the developers of the NC promptly threw away most of the potential advantages by deconstructing the computer in the wrong way.

Most NC designers have so far re-created the capabilities of desktop computers, but with the components in different places. True, they save some money, but they also create vulnerabilities and commercial conundrums. For instance, the NCs now proposed are useless without a connection to the central store—where they can access the software that will make them something other than a silicon-and-plastic paperweight. Larry Ellison of Oracle reckons the low price of an NC will make it a hit in the consumer market. But he glosses over the question

of who will provide a consumer with the necessary disk space and software—and on what terms. Given that fully equipped personal computers, with all the networking support of a basic NC, are expected to cost less than $1,000 by 1997, the price advantage of an NC would be rapidly eroded by even modest software and disk-rental charges.

IBM, predictably, has its eyes on the corporate market for NCs. Corporations are less sensitive to purchase price than to management costs: they must upgrade hundreds of copies of software and back up tens of thousands of files. And companies can afford to build big central disk stores that mere consumers can't. But they had better build carefully. Centralizing essential software means that when the central store breaks, all work comes to a halt. It also means that networks must be carefully designed to accommodate peak loads, lest everything stop just when work is busiest.

The NC technology could have been used just as easily, and much more usefully, to divvy up a computer's functions in a different way. The Nokia 9000 digital cellular telephone, due on the market this fall, has some of the capabilities that the NC designers missed. The Nokia 9000 is both a portable telephone and a portable Internet terminal. About the size of a conventional cellular phone but thicker, it opens to reveal a small screen and keyboard, like an electronic pocket organizer. With built-in networking software, it can easily, and wirelessly, receive e-mail or browse Web pages—although, given the screen size, probably only text Web pages.


DAVID SUTER

The snag with the Nokia 9000 and its ilk, though, is getting them to cooperate with other computers. After all, you don't really want your e-mail in two or three different places, and it would be nice if the addresses entered on your last business trip were also available on your desktop computer when you return. The plug-and-play intelligence offered by Java could easily offer much of that convenience—if only designers could have combined the best of the NC approach with something as fully deconstructed as the Nokia 9000.

Instead of using the network to unify the components that make up a single computer, why not use it to enable a variety of semi-independent computing devices to work together, peer to peer, to create wholes greater than the sums of their parts? How about, for example, a big flat screen, pen-sensitive but without keyboard? At the desktop it doubles as conventional screen and graphics tablet. But take it on the road and it has only enough intelligence and software built in to remember scribbled notes and to surf the Web (by plugging into, say, the Nokia 9000).

Equally, in tomorrow's modern home, the television set and computer might wish to swap information. Thus, Timmy can read more on the rare meerkats he saw on last night's nature show at the Web site whose address the producers kindly intercast—that is, embedded in the broadcast. And at the office, new network-supplied computing services might be made to provide bursts of specialized processing power—for example, the number crunching that is needed to run a simulation.

Java, together with a bit of (preferably wireless) networking, can make most of this integration happen. Indeed, the Federal Communications Commission recently set aside spectrum for exactly the kind of wireless local-area networks needed. But, for the most part, companies have not yet grasped the possibilities. Oracle and IBM sell the hardware and software that run the services on which NCs depend—so they are unlikely to be the first to jump to a vision of networked computing that makes central services unnecessary.

Sun, however, is quietly licensing Java to work with everything that plugs in. Recently it signed a deal with Nortel, a Canadian telecommunications giant, to build Java into cellular-phones-cum-computers. Maybe truly networked computers aren't that far off after all—even if they're not the ones now grabbing the headlines.

—John Browning in London

COMPUTING

Recently Netted....

Debunking Bad Anthropology. Irked by the half-baked anthropology on view at many World Wide Web sites, Candice Bradley, a cultural anthropologist at Lawrence University, started her own page: Classics of Out(land)ish Anthropology (http://www.lawrence.edu/dept/anthropology/classics.html). On it she lampoons the scientific solecisms that catch her eye, from news of Bigfoot to a Web site offering "evidence"—aired on NBC—that humans lived at the time of the dinosaurs ("Human footprints found side-by-side with dinosaur tracks"). One of her favorite targets is the Project Candide Web site; it contains the saga of a trip to Tanzania and Kenya that begins with the voyagers having pizza for the "last time" before departing the U.S. "This is typical of the biased representations of Africa on the Web," Bradley says. "In fact, there are more good restaurants in Nairobi than in most U.S. cities. Pizza is abundant." She also remarks on the safari's maps. They "are tinted with a sienna background so that they resemble 17th- or 18th-century maps of Africa. They are classic examples of the nostalgia for precolonialism and colonialism so prevalent in depictions of things African."

Faxing to E-mail. Two years ago Jaye Muller, a German-born rapper and rock singer, searched for a handy way to receive his faxes with his e-mail. No such service existed, so he created one. Now 24 years old and living in New York City, Muller is president of a company that links personal fax numbers to e-mail addresses—fax to the number, and the document, including graphics and signatures, will appear in the recipient's e-mail in-box. Received as a compressed graphic bit-map file, the fax arrives as a MIME-encoded e-mail attachment. (MIME is the emerging Internet standard for binary attachments to e-mail—that is, for attachments other than plain text, which is sent in ASCII.) His company plans to add voice-mail messages that will also arrive as MIME-encoded e-mail attachments, neatly providing a unified service for fax, voice and e-mail transmission (see http://www.jfax.net). *—Anne Eisenberg (aeisenberg@duke.poly.edu)*

MEDICAL TECHNOLOGY

MAGNET ON THE BRAIN

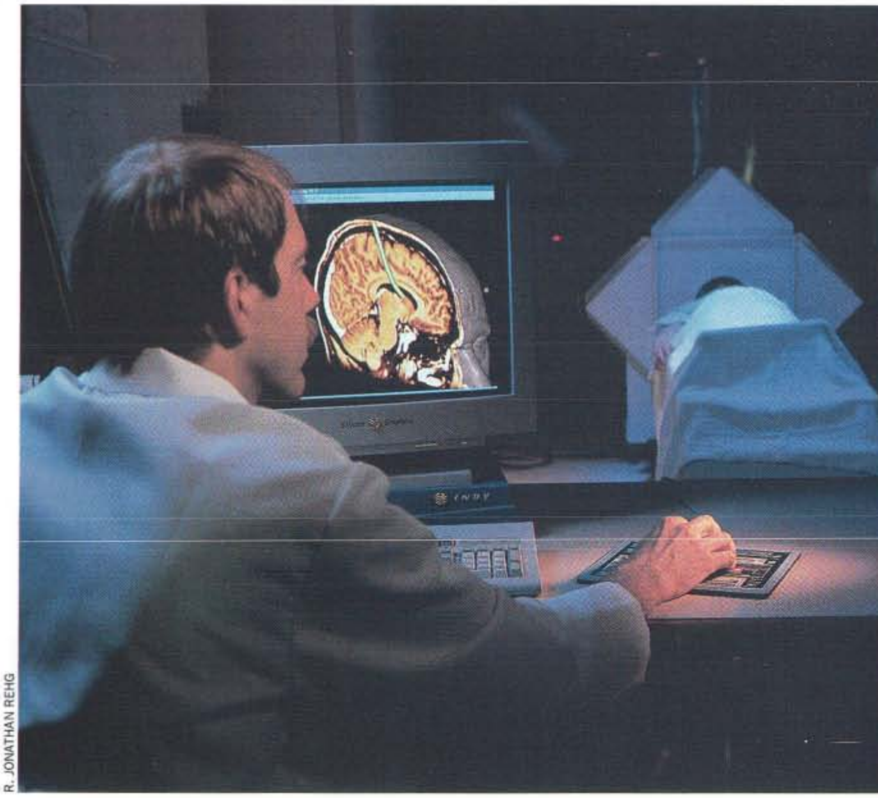
Safer neurosurgery with magnetically steered implants

Children often learn about magnetism by dragging a paper clip through a paper maze with the aid of a magnet held below. Researchers now hope that before long they will accomplish a similar feat in the maze of the human brain with a refined version of a procedure called stereotaxis. The technique, being tested by workers at Stereotaxis, a firm in St. Louis, and at the Washington University School of Medicine, would allow physicians to reach diseased areas of the brain with the least possible damage to healthy tissue.

Stereotaxis is the procedure in which surgeons plunge, say, needles or electrodes straight through the brain to treat a trouble spot deep within. In the process, they tear healthy and perhaps essential neural tissue—a risk complicated if several needles or electrodes need to be inserted, as is sometimes the case. (For instance, to treat Parkinson's disease stereotaxically, six drug-delivering needles would be inserted in different spots to saturate fully the deep-seated striatum, which contains the defective tissue.) Physicians try to minimize surgical damage by first reviewing a brain-scan image and then avoiding the most crucial areas.

The magnetic version of stereotaxis is in principle less destructive. Surgeons would insert a magnetic pellet the size of a rice grain into a small hole drilled into the skull of a patient. The patient's head would then be placed in a housing the size of a small washing machine, which contains six superconducting magnets. Using a magnetic resonance image as a guide, surgeons would then direct the pellet through the brain by adjusting the forces of the various magnets. The pellet could tow a catheter, electrode or other device to minister to the troublesome neural tissue.

With magnetic steering, surgeons can dodge especially critical neurons. Moreover, they would also be able to move the pellet around within the entire damaged area. A patient being treated for Parkinson's would, therefore, have only one path of neurons damaged, as opposed to six with the conventional method.



R. JONATHAN REHG

MAGNETIC STEREOTAXIS SYSTEM would be controlled remotely via a computer. As demonstrated here, a surgeon would use a preoperative brain image to steer a neural implant. The patient's head would lie in a housing that contains the magnets.

The chief obstacle to applying this technique in the past, notes Ralph G. Dacey, Jr., of Washington University, who directs the stereotaxis research team, has been accurately controlling the magnetic fields. A decade ago, however, Matthew A. Howard III, then a physics student at the University of Virginia, realized that the precise instruments physicists use to measure gravity could be applied to the control of magnetic fields. That recognition, coupled with improved computers and brain-imaging devices, enabled investigators to fashion the magnetic stereotaxis system, explains Howard, now a neurosurgeon who assists the researchers in St. Louis from his base at the University of Iowa.

The team has demonstrated the technique on brains from dead mammals and one from a live pig, as well as on a block of gelatin, which has about the same consistency as the human brain. For the moment, other neurosurgeons remain cautious about the system's prospects, and Stereotaxis, which holds the patent on the technique, is the only company committed to this kind of magnetic neurosurgery.

Howard says that although the hardware for magnetic stereotaxis will probably cost more than the conventional technology, it might nonetheless save money by reducing operating time by one half to two thirds. The technology could also be broadened to include use in other parts of the body, such as the liver or blood vessels.

"The challenge," Dacey remarks, "is to find the best complementary use of conventional stereotaxic surgery and specific situations for magnetic stereotaxis." He plans to apply to the Food and Drug Administration before summer's end for approval to start tests with the new method on humans. The first clinical trials, probably for biopsies, could begin next year. *—Philip Yam*

PATENTS

TOO MUCH FOR TOO LITTLE

The Patent Office is swamped with gene sequences it can't afford to check

Every gene sequence that the U.S. Patent and Trademark Office receives must be checked for novelty and obviousness. The PTO uses two massive parallel-processing computers that compare the sequences against five databases; this electronic search is then evaluated by an examiner and, often, a senior examiner. All fairly straightforward.

The problem is that to do this the PTO needs tens of millions of dollars and 100 years—and that's just to review the pending patents. According to John Doll, head of the PTO group that handles gene patents, it takes about 65 hours and \$5,000 to examine a batch of 100 sequences. But the application fee is only about \$800, and some applicants, including Incyte Pharmaceuticals in Palo Alto, Calif., and Human Genome Sciences in Rockville, Md., submit thousands of sequences in an application. As equipment becomes more powerful and automated analysis enables sequences to be tested more quickly for potential pharmaceutical uses, even more applications will be submitted.

In April the PTO held hearings on this crisis in La Jolla, Calif., and Arlington, Va. Commissioner Bruce A. Lehman and attendees suggested possible remedies: raising application fees, bringing in additional examiners or seeking assistance from other agencies. Industry representatives testified that part of the difficulty is that the PTO is doing excessive sequence analysis and that its databases have redundant sequences that slow down analysis.

A similar muddle is slowing down another division of the PTO as well. When a software idea is submitted, it has to be compared with more than a million "prior art" items from the past 30 years. (Prior art is any earlier patent, journal article, book or news story that anticipates the invention.) Examiners have yet to be provided access to the databases and tools they need. The Patent Office seems to be suffering from too much of a new thing. —*Gregory Aharonian*

IMAGING TECHNOLOGY

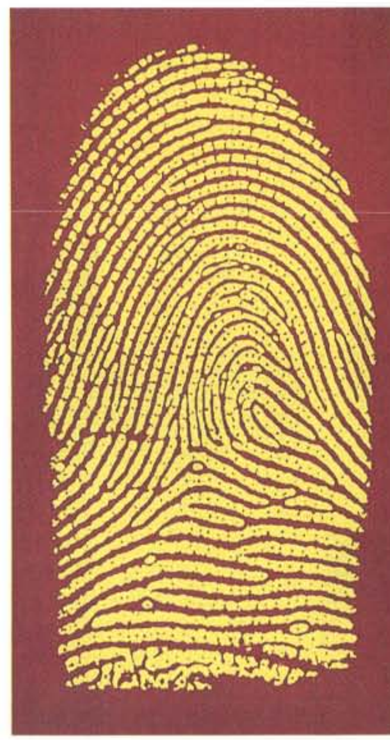
The Right Touch

With just a tap of your finger, imagine unlocking your house, withdrawing money from your bank account or even shopping. It may seem like a futurist fantasy, but electronic fingerprint identification can no longer be relegated to the realm of science fiction.

Although the technology has been in the works for a few years—and New York City–area airports have used it on a limited scale since 1994—it is finally becoming widespread. This past April the New York City police department contracted with two companies specializing in biometric security systems—MORPHO Systems and Identix—to install a fingerprint identification system. Now an officer will be able to scan a suspect's fingerprints into a database, take a digitized mug shot, type in other details and then electronically send the entire package to headquarters. The system is expected to be more accurate than current paperwork procedures, sparing police thousands of hours.

Other government agencies have already jumped on the electronic identification bandwagon, and many state social services departments, including those in New York, New Jersey and Connecticut, have such systems for identifying welfare recipients. (To date, for instance, New York's Suffolk County has documented saving more than \$1 million, mostly by blocking false claims.)

The technology works by photographing the swirls and whorls of each fingertip. A computer tabulates and records the locations of specific ridges, indentations and patterns known to be unique to each person. Identix reports that its scanning equipment is nearly 100 percent effective in matching the right person with the right fingerprint—but the computer has also rejected a correct match 3 percent of the time. (The chance of any two people having the identical fingerprint is estimated to be less than one in a billion.) If the technology keeps reaching wider and wider audiences, you, too, may soon be asked for your hand. —*Gunjan Sinha*



SCOTT CAMAZINE *Photo Researchers, Inc.*

SEMICONDUCTORS

ONE SMALL STEP

The next big advance in chip design arrives one year early

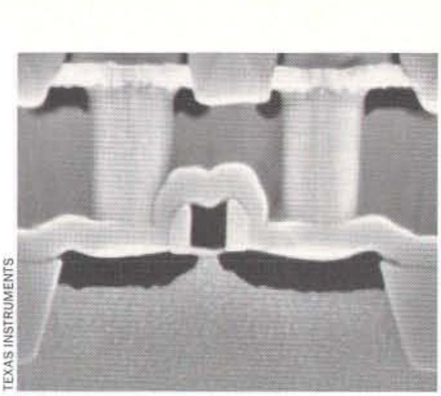
With grand fanfare, the electronics giant Texas Instruments announced in May that it had perfected a process that can produce silicon microchips of far greater detail and complexity than any currently available. Newspapers widely marveled at the innovation; many pointed out that TI is the first to produce chips with features as small as 0.18 micron (millionths of a meter) wide. Some predicted that the microchips would launch a generation of wonderfully smart and compact contraptions.

Such reports were wrong on two counts, but correct on the third. TI was not first. Although that company has prototypes on hand and hopes to have a factory constructed by next year, IBM began shipping small quantities of equally detailed integrated circuits in May. And both TI's and IBM's processes create tiny transistors that are 0.25, not 0.18, micron in width. (The much mis-

understood 0.18-micron measurement refers not to feature size but to the distance current must travel to switch a single transistor.) This long-expected advance is the logical next step beyond the 0.35-micron features that make up the Pentium Pro and PowerPC chips now on the market, but it was not anticipated to occur until 1997.

Hyperbole aside, the new semiconductors may indeed have a dramatic impact on computers over the next five years or so, for several reasons. First is power consumption. TI claims its devices run on as little as one volt—about one third the voltage required by Intel's Pentium. Such low-power chips could significantly extend battery life in portable gadgets. The second benefit is sheer size.

Whereas a Pentium Pro today spreads about 3.3 million transistors across four layers, the new processes draw smaller switches onto six layers, with even more layers to come in the near future. TI says it can pack up to 125 million transistors onto each new chip—but that is true only if there are no wires connecting them. A more realistic estimate is about 20 million. Whether it can do so without also charging four times more for its chips than Intel does for the Pentium Pro remains to be seen.



TEXAS INSTRUMENTS

TINY TRANSISTOR, down to 0.25 micron in size, represents the next generation of chip.

Greater breadth and depth lead directly to the final advantage: speed. Smaller transistors always switch faster, but the real boost will come from combining into a single chip functions that previously required several processors. "A lot of speed is lost when you have to move signals between chips," observes G. Dan Hutcheson, an industry consultant at VLSI Research. In particular, he suggests, chip designers will want to combine memory with logic circuits in ways they never could before.

For all its promise, the advance also represents a threat to the computer industry. "As we push below 0.25 micron, the software tools available to design integrated circuits are not going to be able to keep up with the added complexity," Hutcheson warns. If manufacturers have to add dozens of engineers to produce each new design, chips will not remain cheap—and fast-evolving—for long. To head off what it calls a "productivity gap," the industry consortium SEMATECH awarded a multimillion-dollar contract to Synopsys in May for an advanced design system that can handle circuits of 0.25 micron and the next two or three smaller increments. Beyond that lie limits that will force chipmakers to look for great technological leaps rather than small, safe steps. [See "Technology and Economics in the Semiconductor Industry," by G. Dan Hutcheson and Jerry D. Hutcheson; SCIENTIFIC AMERICAN, January.]

—*W. Wayt Gibbs in San Francisco*

MATERIALS SCIENCE

Coat of Many Colors

Conventional wisdom has it that red cars attract more speeding tickets. But what about cars that change color? Several new coatings may soon permit drivers to test their legal luck.

Taking cues from nature, chemists have been able to develop paints that derive their colors from interference patterns. The brilliant colors of butterflies, for example, result from multiple layers of extraordinarily thin fibers found in the insects' wings. When light falls on the wings, the top layers reflect the rays at a slightly different angle than the bottom layers do. The different reflected wavelengths then interfere with one another, producing new wavelengths that appear as shimmering colors.

MERCEDES-BENZ AG

The use of such coatings has been limited to small objects, until now. Several companies have recently described their efforts to create car paint based on this principle. Researchers at Nissan and the Tokyo Institute of Technology spun tiny strands of polyester that gave rise to interference patterns—and the iridescent blue seen in certain butterflies. Mercedes-Benz is offering European customers paint that changes color depending on one's viewing point: light reflects off layers of liquid-crystal polymers at different angles, producing various colors. And Ford offers a limited-run 1996 Mustang with paint that can appear green, purple, gold or amber.

RICHARD MIKALA *Bruce Coleman, Inc.*

Not surprisingly, these unusual paint jobs remain a luxury option: Mercedes charges 10,000 deutsche marks (that's U.S.$6,600) for the customized work. The vacuum technology that is needed to produce the paint is very expensive, and the coatings themselves can be difficult to handle because the microstructures that produce the colorful interference patterns can break, particularly in the application process. No word yet on how well they tolerate fender benders. —*Sasha Nemecek*

PROFILE: SHELLEY A. HARRISON

Exploring the Business of Space

The curving tower of yellow smoke that just lofted *Endeavour* into the morning sky over Cape Canaveral on Florida's east coast is beginning to disperse, and Shelley A. Harrison, whose company has entrusted much of its assets to the space shuttle's cargo bay, is beaming. Although he missed a night's sleep schmoozing and talking business, Harrison, chairman and CEO of Spacehab, exudes confidence. The commercial space business is poised to take off, he believes, and Spacehab has—for now—no competition. "I believe human habitation of space is going to happen, and Spacehab's objective is to support it," Harrison declares.

Harrison, possibly the only person at the Kennedy Space Center wearing a suit and tie, is a high-tech venture capitalist with a mission to commercialize the space frontier. His voice is academic-precise, rather than big-business-brash, harking back to his days as a university scientist. But Harrison has the kind of record that commands attention in the world of commerce: one of his early ventures helped to establish the bar codes that now adorn products throughout the developed world. Harrison thinks low earth orbit is a territory as ripe for technological development as the retail stores of the 1970s.

Although few companies to date have invested much effort in space-based research—and almost none are contemplating manufacturing in space—Harrison believes that will change. As he sees it, opportunities to put payloads in orbit on the shuttle have been too infrequent for most businesses to evaluate the idea seriously. Spacehab aims to jump-start orbital industry by providing customers with room in laboratory modules that fit in the shuttle's cargo bay. The company, which leases space on shuttles from the National Aeronautics and Space Administration, has won a contract to carry supplies for the Russian *Mir* space station into orbit. Harrison has an even more ambitious long-term target: privatizing operations on the planned International Space Station.

Harrison's parents, who named him after the poet, hoped he would become a rabbi. He studied the Talmud and says the logic and argumentation taught him "to think on multiple levels." But it was the post-Sputnik era, and so after a year studying physics in Israel, Harrison decided to accept a NASA scholarship that enabled him to earn a bachelor's degree in electrical engineering. He went on to work on military phased-array radars at AT&T Bell Laboratories.

While working in the 1960s toward a doctorate at Brooklyn Polytechnic (now Polytechnic University), Harrison had the opportunity to observe at close quarters the firm Quantronix, which was struggling, without much success, to develop lasers for exotic applications. The experience taught him an important lesson about business. "I realized that were I ever to do that—go and form a company that might involve lasers—I'd look for a very pervasive, wide market application, not little niche markets."


GREGORY HEISLER

Ph.D. in hand, Harrison became a professor at the State University of New York at Stony Brook, where he was in charge of developing the quantum electronics curriculum. He learned the art of grantsmanship and also formed a nonprofit concern, Public Systems Research, to allow students and faculty to supplement their incomes through consulting. Among the clients were NASA, which was designing combustion experiments for the Skylab missions flown in the 1970s, and the Universal Product Code Council, an organization that was seeking a machine-readable way to represent product information.

By the early 1970s, supermarket clerks were entering an increasing amount of data into machines, but they were still doing it by hand and thus making many errors. Lasers, Harrison realized, could provide "eyes for the computer" in an automated system. He teamed up with Jerome Swartz to establish Symbol Technologies and then to "conquer the world," as Harrison puts it.

Symbol's goal was to develop portable laser bar-code scanners, but first there had to be bar codes to scan. Harrison and Swartz, joined later by Harri-

son's wife, Susanne, started (in, yes, a garage) using a computer-driven device to print bar codes onto film and incorporating the result into product packaging. Symbol introduced a handheld laser scanner attached to a fixed station that printers and packagers could use to check the readability of bar codes.

Symbol then automated the magazine-returns industry with an expanded code that made issue dates machine readable, and the inventors later perfected more portable "gun scanners" that have since invaded record and bookstores and even department stores. (They avoided supermarket checkout desks, Harrison explains, because it became clear that cash-register manufacturers were better positioned to corner that market.)

Symbol made investors "a ton of money" after it went public in 1979, Harrison says. But he quickly realized he was more interested in blazing new high-technology trails than managing a fast-growing company. In 1982 he and Susanne formed their own venture-capital business, Harrison Enterprises. In the late 1980s, together with Herman Fialkov and others, Harrison created Poly Ventures, which raised \$53 million to invest in start-up semiconductor, laser and software companies, several of which are now publicly traded. (Under a profit-sharing arrangement, \$1.5 million of his profits to date have gone to his alma mater, Polytechnic University.)

"Is [space commerce] evolving into something better? Yes, and I think Spacehab is doing it."

Meanwhile Harrison had formed a close relationship with Robert Citron, a "visionary, explorer type" who had been an international projects administrator at the Smithsonian Institution and then went on to found Spacehab. Citron asked Harrison to help him boldly go where no business had gone before. Citron's original vision of space tourism on the shuttle soon evolved into a plan to provide laboratory space. Whereas most aerospace companies content themselves with bidding for government contracts to supply hardware, Spacehab was founded on the principle that it would own its hardware and lease to the government.

Because NASA officials originally envisioned the shuttle as merely a vehicle for carrying cargo to a space station, it had little room for experiments. But when it became clear the station would be long delayed, Harrison and Citron began playing with the notion of doing research on laboratory modules in the capacious cargo bay. Astronauts could gain access to a module through a tunnel connected to the shuttle's mid-deck. Citron reckoned that by avoiding unnecessary bureaucracy, a private company could provide modules for less money than NASA could.

Harrison raised the idea with aerospace executives, who approved of the concept but declined to invest. In 1987 Citron convinced Harrison, who had already put some of his own money in Spacehab and had a seat on the board, to devote himself to raising capital. It was by now clear that the first three laboratory modules—the minimum number worth building—would cost more than \$120 million. Harrison found believers in Europe, where Daimler-Benz and Alenia Spazio are investors, and in the Far East, where Mitsubishi, wealthy investors in Taiwan and the government of Singapore's venture arm have sizable stakes.

But the entrepreneur was still short of the target when a NASA-commissioned study concluded that the agency would have to spend \$1.2 billion to build flight hardware with the capabilities that Spacehab was offering for \$185 million. Spacehab got the contract, bridged the capital gap and started building. Harrison became chairman in 1993, the same year as Spacehab's first flight. *Endeavour*'s mission this past May was the fourth for a Spacehab laboratory module and the fifth for the company (a storage module carried supplies to *Mir* earlier this year).

Spacehab has increased annual profits since its initial flight, although it is still showing a cumulative loss since its inception. But if Harrison's vision is borne out—and if the U.S. keeps the shuttle fleet in service—the company could reap large returns. Harrison is now negotiating places for Spacehab modules on NASA's existing space shuttle manifest. A little shuffling around of payloads in the shuttle, he maintains, can free up room that can be sold to companies as well as NASA's partners in the International Space Station for testing station equipment. Spacehab has recently introduced a new double module that extends the possibilities, he points out.

Harrison's two children seem to have acquired from their father an enthusiasm for technology. Rachel is a systems engineer and artist who is now developing interactive multimedia, and Daniel is training to be an electrical engineer. The senior Harrison scorns those academic scientists who sneer at space-based research but have never put money at risk. "If I'd listened to all the naysayers about the prospects for turning technology into profit with respect to my ventures, I wouldn't have done any of them," he grumbles. He trusts his "gut" to sense opportunities.

Harrison maintains that although commercial space research has been conducted sporadically in the past decade, it got under way "in earnest" only with Spacehab's first flight. So it is impressive, he argues, that in 1995 industry and academia put up \$38 million in cash and "in-kind contributions"—which include materials—for commercial space research. NASA put up about half that amount and provided free launches, but Harrison defends the subsidy as appropriate for a fledgling industry.

He sees grounds for optimism in the 100 companies on the books as commercial space affiliates of NASA as well as the five new space-related patents granted and 11 filed in 1995. "The commercial microgravity research conducted to date clearly has demonstrated that there is value," he states, noting its "iterative nature." He points to Kennametal, an industrial tool manufacturer that has flown four experiments on Spacehab modules to learn more about liquid-phase sintering of alloys, as well as the orbital activities of some pharmaceutical companies that may lead to new drugs. Where many observers find fragmentary results and tepid interest, Harrison sees the start of a trend.

Harrison's hunch is that industry will embrace experiments in space—and possibly even manufacturing—when it can negotiate with a commercial partner and be sure that standardized orbital facilities will be available. "The way things work now, is it viable to carry out commercial activity and research in space for some benefit on the earth? Hardly likely," he concedes. Now the venture capitalist is warming to his subject, and he raises his voice just a little. "Is it evolving into something better? Yes, and I think Spacehab is doing it. Can it be good and worthwhile? Yes. We believe it, or we wouldn't be putting our money and the company on the line to develop it."

—Tim Beardsley at Cape Canaveral

Smart Cards

As potential applications grow, computers in the wallet are making unobtrusive inroads

by Carol H. Fancher

The semiconductor revolution has advanced to the point where the computing power that once took up an entire room can now be lost among the spare change, house keys or candy wrappers in the average pocket. For more than 10 years, "smart" credit cards incorporating tiny chips have been in use in France and other parts of Europe. A set of standardized contacts on the front of each card supplants or supplements the familiar

ATLANTA, GEORGIA, is the site of the largest trial thus far of smart cards in the U.S. More than one million cards will be sold in conjunction with the 1996 Olympic Games. Cards can be used at Olympic event sites and at restaurants and shops throughout

coded magnetic stripe on the back. Although the U.S. has been lagging in the use of this technology, a series of ongoing pilot programs may soon change that situation. Some pundits have criticized smart cards as a technology endlessly in search of meaningful applications, but the divergent experiences of different countries show that the issues are more complicated.

Curiously, telecommunications policy has been one of the major influences on the deployment of smart cards. In the U.S., where telephone calls are cheap and it is a simple matter to attach a magnetic-stripe reader to a phone line, the fraud-reduction aspects of smart cards are not necessarily worth the extra expense. Instead merchants can dial up a central database to make sure a card is valid before completing a transaction. In Europe, where calls are generally more expensive and connecting modem-equipped devices to phone lines is more difficult, security was a significant driving force behind smart-card introduction.

PHOTOGRAPHS BY ERIK S. LESSER

the city. Turnstiles in the subway system also accept the cards.

The French, for example, made the switch during the mid-1980s because fraud rates were unacceptably high and rising. With smart cards, merchants do not have to go on-line to centralized databases. They can rely on personal identification numbers (PINs) to verify the ownership of a card simply by checking the PIN typed in by a customer against the record on the card itself. Furthermore, the chips are more resistant to tampering than magnetic stripes, which can be read and written on with readily available equipment. Over 20 million smart cards are now in use in France.

One motivation for smart-card introduction in the U.S. today is the possibility of multiple uses for the same card. In theory, the same silicon-imbued piece of plastic could serve as personal identification, credit card, automated teller machine (ATM) card, telephone credit card, transit pass, carrier of crucial medical information and cash substitute for small transactions in person or over the Internet. Additional uses are limited mostly by issuers' imaginations and consumer acceptance. As a single card becomes able to hold more parts of a person's life, security and privacy concerns will have to be met; cards of the future will probably be highly personalized.

Standardizing Intelligent Transactions

Smart cards are becoming more attractive as the price of microcomputing power and storage continues to drop. They have two main advantages over magnetic-stripe cards. First, they can carry 10 or even 100 times as much information—and hold it much more robustly. Second, they can execute complex tasks in conjunction with a terminal. For

BETH PHILLIPS

SMART CARD contains memory and a micropressor underneath gold contact pads. The position of the pads is governed by standards so that cards and readers from many sources can work together.

example, a smart card can engage in a sequence of questions and answers that verifies the validity of information stored on the card and the identity of the card-reading terminal. A card using such an algorithm might be able to convince a local terminal that its owner had enough money to pay for a transaction without revealing the actual balance or the account number. Depending on the importance of the information involved, security might rely on a personal identification number such as those used with automated teller machines, a midrange encipherment system, such as the Data Encryption Standard (DES), or a highly secure public-key scheme.

Smart cards are not a new phenomenon. They have been in development since the late 1970s and have found major applications in Europe, with more than a quarter of a billion cards made so far. The vast majority of chips have gone into prepaid, disposable telephone cards, but even so the experience gained has reduced manufacturing costs, improved reliability and proved the viability of smart cards. International and national standards for smart cards are well under development to ensure that cards, readers and the software for the many

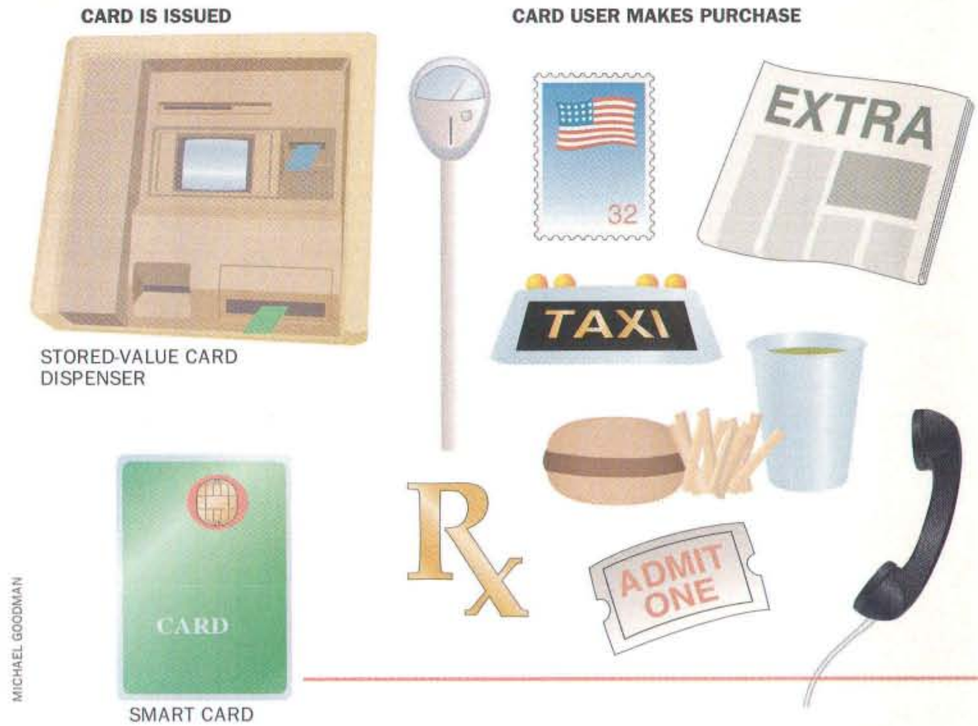
different applications that may reside on them can work together seamlessly and securely. Standards set by the International Organization for Standardization (ISO), for example, govern the placement of contacts on the face of a smart card so that any card and reader will be able to connect.

Industry-specific standards are being developed for cards to be used in applications as diverse as digital cellular phones, satellite and cable television and, of course, finance. Recently Visa, MasterCard and Europay agreed on a common specification for smart cards that defines the basic protocols for communication between cards and readers (analogous to the RS-232 standards that govern communication between personal computers and modems). The specification is general enough so that virtually any kind of information can be exchanged by hardware and software that conform to it. As a result, this agreement could bring the convenience of a single card for purchases, ATM withdrawals, frequent flier miles and even Internet access.

Under the Hood

Standards dictate a card's shape and electrical connections, but the technology inside has gone through significant evolution. The simplest "memory" cards contain only nonvolatile memory and a limited amount of logic circuitry for control and security. They typically serve as prepaid telephone cards—a terminal inside the pay phone writes a declining balance into the card's memory as the call progresses; the card is discarded when its balance runs out.

Smart cards are more sophisticated and contain a chip with a central processing unit and various kinds of short- and long-term memory cells. Some versions may also incorporate a special coprocessing circuit for cryptographic operations to speed the job of encoding and decoding messages or generating digital signatures to validate the information transferred. [For more information on the kinds of cryptographic protocols that could be employed in smart cards, see "Confidential Communication on the Internet," by Thomas Beth; SCIENTIFIC AMERICAN, December 1995, and "Achieving Electronic Privacy," by David Chaum; SCIENTIFIC AMERICAN, August 1992.] Smart-card standards place no limitation on the amount of processing power in the card as long as the chip in question can fit the space



STORED-VALUE CARDS are electronic analogues of the traveler's check. They can be used to purchase items ranging from fast food to parking. Consumers buy cards already loaded with monetary value from a dispenser and use the cards for small trans-

allotted for it under the contact pad.

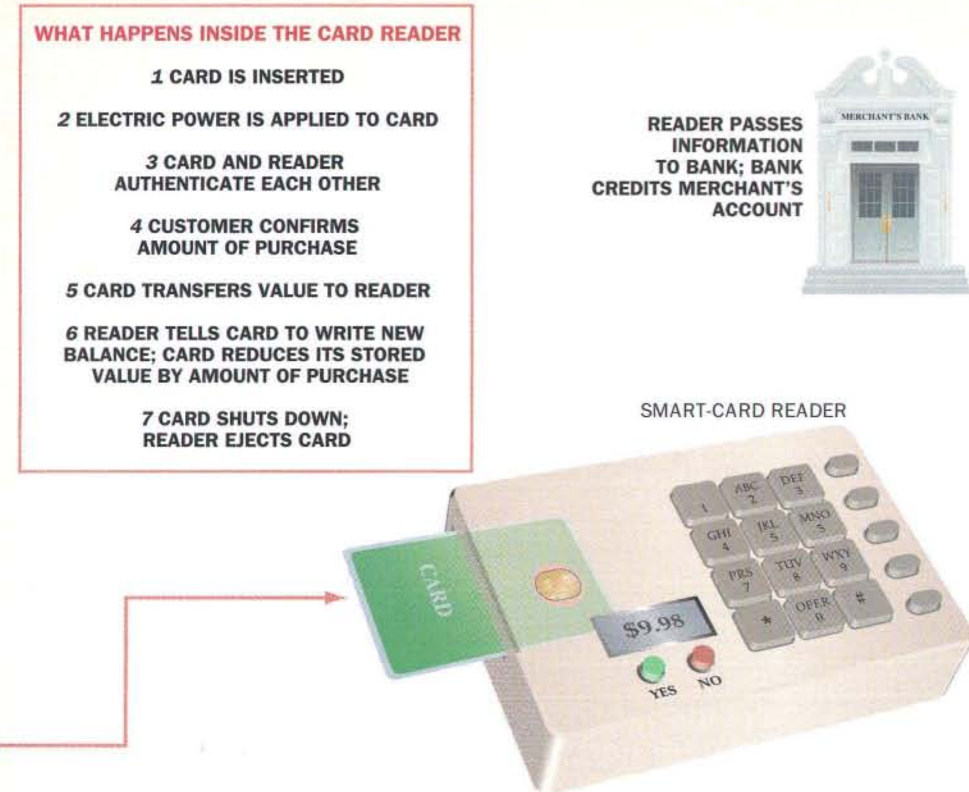
Current smart cards, made by firms such as Giesecke & Devrient, Gemplus, Schlumberger and Solaic, range in price from less than $1 to about $20. (The silicon inside the cards is made by companies such as Motorola, Siemens and SGS-Thompson.) A magnetic-stripe card, in contrast, may cost between 10 and 50 cents, depending on whether the card is bare or incorporates a photograph or a holographic patch and on how many cards are made at once.

Because the cards are dependent on an outside power source provided by the reader interface, any information held in conventional random-access memory (RAM) will be lost every time it is removed from a reader. Hence, smart-card microprocessors use only a few hundred bytes of RAM as a scratchpad for working on transactions in progress. The software that controls a card's operations must survive from one use to the next, and so it occupies between three and 20 kilobytes of permanent nonvolatile read-only memory (ROM). The contents of the ROM are fixed in the chip when it is made. The personal, financial or medical data that give each card value to its owner reside in an alterable nonvolatile memory (EEPROM, for electrically erasable programmable read-only memory) of between one and 16 kilobytes.

The need for security influences the design and handling of the card, its embedded circuitry and its software. Microprocessors used in smart cards are specifically designed to restrict access to stored information and to prevent the card from use by unauthorized parties. Typically a card will work only in a well-characterized operating environment.

For example, criminals may attempt to force the card to operate outside certain voltage or clock frequency ranges in the hope that it will display weaknesses that can be exploited; a properly designed device will automatically fail to respond under such conditions. In some cases, circuit links may be designed to become inoperable once a card has been programmed, so that vital data cannot be altered. Manufacturers also employ special tamper-resistant techniques that would prevent a thief from getting to the microscopic circuitry directly.

Most smart cards require physical contact between the card and pins in the reader, but a growing set of applications depends on so-called contactless cards. Short-range cards operate by electrical inductive or capacitive coupling with the reader and card a millimeter or so apart;



actions. Card readers transfer information to banks periodically for credit to the merchant's account, either directly or through a clearinghouse. Sophisticated stored-value cards may be reloaded; simple ones are discarded when their cash is used up.

longer-range ones communicate by radio signals. (The radio-frequency energy emitted by the reader also powers the cards, which must therefore be extremely sparing of current.) Contactless smart cards are often used in situations where transactions must be processed very fast, as in mass-transit turnstiles. Transit system operators in Hong Kong, Washington, D.C., Manchester, England, and about a dozen other cities have tested contactless cards; Hong Kong will issue three million cards by 1997.

Developers and users are working together to develop firm standards for long-range contactless cards. Efforts are also under way to standardize hybrid cards that can communicate either directly or by radio links. Lufthansa, the German national airline, has already begun issuing a hybrid card to frequent fliers; the contactless part serves as an ID card for the firm's paperless ticketing system, and the contacts make for a European-standard smart credit card. Roughly 350,000 will be in circulation by year's end.

The smart card is a technical achievement in its own right; it is, however, merely the most identifiable part of a vastly larger transaction system that surrounds it. The traits of this infrastructure may have much more influence on the evolution of the card's role in society than do the characteristics of the card itself. It is therefore important to see how the card would function as part of the larger system to understand why it might be appealing.

The Big Picture

Consider, for example, the stored-value card, at present the most common application of chip-card technology. The attractions of such a card hinge on the relatively high overhead costs of alternatives such as credit cards or cash. Even in the U.S., verification costs are too high to allow a profit on conventional card transactions smaller than a few dollars. The stored-value card minimizes transaction costs by carrying monetary value directly, instead of merely acting as a pointer to an account. It transfers the digital equivalent of bills or coins to a merchant's digital "cash register," whereupon they can be deposited in a bank. Children, tourists and others who do not have a local bank account can use these cards, which can even be sold from vending machines.

Such cards are particularly attractive for pay phones, parking meters, photocopiers and vending machines. By eliminating the coin box, they remove a tempting target for thieves and vandals. Although digital tills must be secured against both unauthorized emptying and stuffing with counterfeit electronic cash, these problems appear easier to handle than their physical counterparts.

Bypassing the handling of money in paper or metallic form could generate significant savings. Economists estimate that counting, moving, storing and safeguarding cash cost about 4 percent of the value of all transactions. The interest lost by holding cash instead of keeping money on deposit is also substantial. The Royal Bank of Canada, which is participating in digital-cash trials in Ontario, keeps about a billion dollars on hand at all times.

The costs per transaction of stored-value cards tend to be lower than those for credit cards and cash, but initial capital costs tend to be higher. The cards themselves cost more, and whoever pioneers their use must bear the expense of installing an infrastructure of card

MONDEX

SWINDON, ENGLAND, is the site of an ongoing trial of Mondex, an "electronic purse" system in which smart cards exchange digital funds. Unlike most other stored-value systems, Mondex allows electronic currency to pass from hand to hand indefinitely without being redeposited. About a quarter of the people in Swindon use the cards at shops, restaurants, laundries and newsstands. Another trial starts this fall in Guelph, Ontario, where even parking meters will accept cards.

readers. In addition, software designed to process transactions by credit and debit card must be modified to deal with the new form, which more closely resembles a digital traveler's check. A typical smart-card reader costs over $100, roughly comparable to the price of the box that reads a magnetic-stripe card and calls a credit-card company to verify a transaction. There are over 13,000 smart-card readers in the U.S. versus more than five million devices capable of dealing with conventional credit cards.

More than two dozen companies are working on smart-card readers, and prices will no doubt drop with volume production. Nevertheless, the amount of equipment that must be installed is substantial. Outside the U.S., the number of stored-value cards is steadily growing, with major national programs implemented or planned in Australia, Canada, Chile, Colombia, Denmark, Italy, Portugal, Singapore, Spain, Taiwan, the U.K. and elsewhere. Levels of consumer acceptance vary; the cards provide clear potential savings for banks and merchants, but transforming those benefits into incentives for users can be difficult. National banking authorities are also understandably cautious about what is in effect a new method of printing money, with no fixed rules about whose authority guarantees its value.

Most stored-value cards now in use are disposable. Reloadable devices would work the same way for making purchases but would have extra software that would enable a consumer to transfer money to a depleted card. (Encryption or other security techniques would help ensure that a card could be recharged only in a legitimate transaction.) Citibank, Chase Manhattan, Visa and MasterCard are assembling a pilot program for stored-value cards in New York City. The companies will issue reloadable smart cards to approximately 50,000 customers; the cards will also have magnetic stripes for conventional transactions. About 500 stores, restaurants and other merchants will have readers capable of accepting electronic-cash transactions. More than one million stored-value cards are also being issued for the 1996 Olympic Games in Atlanta; they can be used in Olympic venues and at several thousand nearby shops.

A number of groups are backing competing smart-card schemes for stored value. All use essentially the same hardware, but their software differs. Manufacturers of card readers are therefore developing equipment capable of handling multiple protocols. It is not yet clear which system consumers will favor, and each has its own strengths and weaknesses. The stored-value protocols of the New York and Atlanta pilot programs, for example, are relatively simple but limited—for example, there is no provision for rescinding or replacing the value of a card that is lost or stolen. The DigiCash system, which relies on complex cryptographic protocols, is both secure and untraceable but requires more processing power and hence more expensive cards. The British Mondex system, meanwhile, is intended as a full-scale secure cash replacement: electronic money can pass from one user to another indefinitely without being redeposited in a bank. A trial is under way in Swindon in the southwest of England, and another one is beginning in Guelph, Ontario, where even parking meters will accept digital currency.

Some Smart-Card Applications

APPLICATION AND LOCATION	NUMBER OF CARDS	STATUS
MasterCard Cash stored-value card, Canberra	10,000 issued	Started in March 1996. Cards usable at 250 stores
VisaCash stored-value card, Atlanta	More than one million	Cards usable at Olympic sites, transit system and several thousand stores
Proton stored-value card, Belgium, Netherlands, Brazil, Australia	90,000 issued	Full-scale introduction in progress
Social Security ID card, Spain	500,000 issued, seven million by 1997, 40 million by 2001	Full-scale introduction in progress. Card gives access to medical benefits and is verified by stored fingerprint
Citizen ID card, South Korea	1,500 issued	Pilot project. Card includes ID, driver's license, medical insurance and retirement benefits
Health insurance card, Germany	80 million issued	Started in 1994 for identification only
Health information card, European Union	200,000 to be issued starting in 1996	Pilot projects for cards containing only essential information for medical treatment
Contactless transit farecard, Hong Kong	20,000 issued, three million by 1997	Pilot project started in November 1995. System-wide introduction in progress
ID and stored-value card, Washington University, St. Louis	12,500 issued	In use. Cards work in vending machines, laundries and other small-value applications. They also serve as ID cards for access to campus facilities

SOURCE: CardTech/SecurTech '96 Conference Proceedings, Atlanta

LISA BURNETT

Protecting Health

In a mark of the technology's versatility, smart cards can also carry vital medical information. Instead of just indicating that a person has medical insurance, for example, a card can store details of the coverage. It can also provide basic medical information, such as lists of drug sensitivities, current conditions being treated, the name and phone number of a patient's doctor and other information vital in an emergency. An intelligent card that carries only the information most relevant to current treatment can streamline care significantly even as it bypasses the potentially intractable privacy and ownership concerns that would arise if health care administrators attempted to place every patient's complete medical history on a chip for easy portability.

Indeed, simply automating the process of entering a person's name and account number into medical forms can make insurance processing much more efficient. Germany has recently begun to issue to all its citizens chip cards that will carry their basic health insurance information, and France is investigating a similar program. Both countries have thus far decided against storing more sensitive data on chips until legal, ethical and security issues can be ironed out.

In France and Japan, kidney patients

can carry cards that hold their dialysis records and treatment prescriptions. Dialysis patients often need their blood cleansed two or three times a week. Each session involves a particular set of machine parameters and a personalized combination of drugs as well as the use of a kidney dialysis machine. Before the introduction of the smart cards, patients could go only to the local dialysis center where their records were kept, but now they have the geographic mobility most of us take for granted. Security checks built into the cards help to ensure that no one except doctors and other authorized persons can read or update treatment information.

Personal Communication

Because the telecommunications costs involved in verifying credit-card transactions have played a crucial role in the history of smart cards, it is perhaps appropriate that one of the device's most innovative applications is at the heart of a new generation of mobile communications. The Global System of Mobile Communications (GSM) is a technical specification for digital cellular telephones; about 10 million people have GSM phones, and service is available or under development in more than 85 countries. Every GSM handset is designed to accept a smart card that carries information about the telephone number of the card's owner and the suite of services it can access. A Swiss executive traveling to Belgium can just remove the smart card from her GSM unit at home and plug it into a rented or borrowed unit at her destination. When callers dial her number, the switching system will automatically locate the handset with her smart card anywhere in the world and deliver the call to it. In addition, the smart card can encrypt the transmission, preventing the casual eavesdropping possible with other forms of cellular phones.

As with other smart-card applications, the U.S. lags behind many nations in GSM services. There are a few pilot programs in place, but widespread deployment is not expected until 1997. The GSM systems being built in the U.S. operate at a frequency of 1.9 gigahertz instead of the 1.8 gigahertz used elsewhere and employ two competing, incompatible technologies. As a result, handsets may be useless outside their home range. The smart cards that animate them, however, should work anywhere.

Cards That Know You

If smart cards can give identity to an electronic device, will they eventually serve as foolproof credentials for humans as well? Smart cards can carry much more information than the paper or plastic rectangles that are used to constitute drivers' licenses, insurance cards or other kinds of identification. And they can probably carry it more securely.

ID cards often have a picture and signature so that authorities can make sure the bearer matches the card. Smart cards can store a PIN to improve security, but they can also add a catalogue of other biometric identifiers: voiceprints, fingerprints, retina scans, iris scans or dynamic signature patterns. Presented with a card holding a reference pattern of some kind, computers can determine with a remarkable degree of accuracy how well its bearer matches that pattern. Customs authorities in the Netherlands have already tested a system to speed passport checking at the airport for frequent fliers: the person puts a finger on a glass plate, and a video camera captures the fingerprint; a computer then compares the video image with a reference print stored on the smart card. With the template on a smart card, there is no need to connect to a centralized database to confirm a person's identity.

Such matching techniques are as yet imperfect—the smart cards function well, but the algorithms for deriving and comparing the biometric patterns are still imperfect. Furthermore, designers must decide whether they are more interested in rejecting impostors or making sure that legitimate cardholders are always accepted. A card that subjects its owner to the embarrassment of an ID mismatch even once a year is unlikely to find wide acceptance.

This consideration and others suggest that smart cards have reached a first plateau of technological maturity: their capacity is no longer the limiting factor in systems that employ them. Instead their future depends on software design, economics, liability and privacy concerns, consumer acceptance and a host of other political and personal issues.

MICHAEL GOODMAN

CELLULAR TELEPHONES based on the GSM standard are lifeless without a smart card to animate them. The card holds the subscriber's phone number and other account information. It can also perform digital signal processing to encrypt the conversation and foil the eavesdroppers who bedevil users of conventional cellular phones.

The Author

CAROL H. FANCHER has been working at Motorola for the past four years to define and develop the U.S. smart-card market. Before joining Motorola, she held engineering positions at Tracor, Ford Microelectronics and the Fraunhofer Institute for Integrated Circuits in Erlangen. In 1979 Fancher received a B.Sc. in electrical engineering from the University of Texas at Austin.

Further Reading

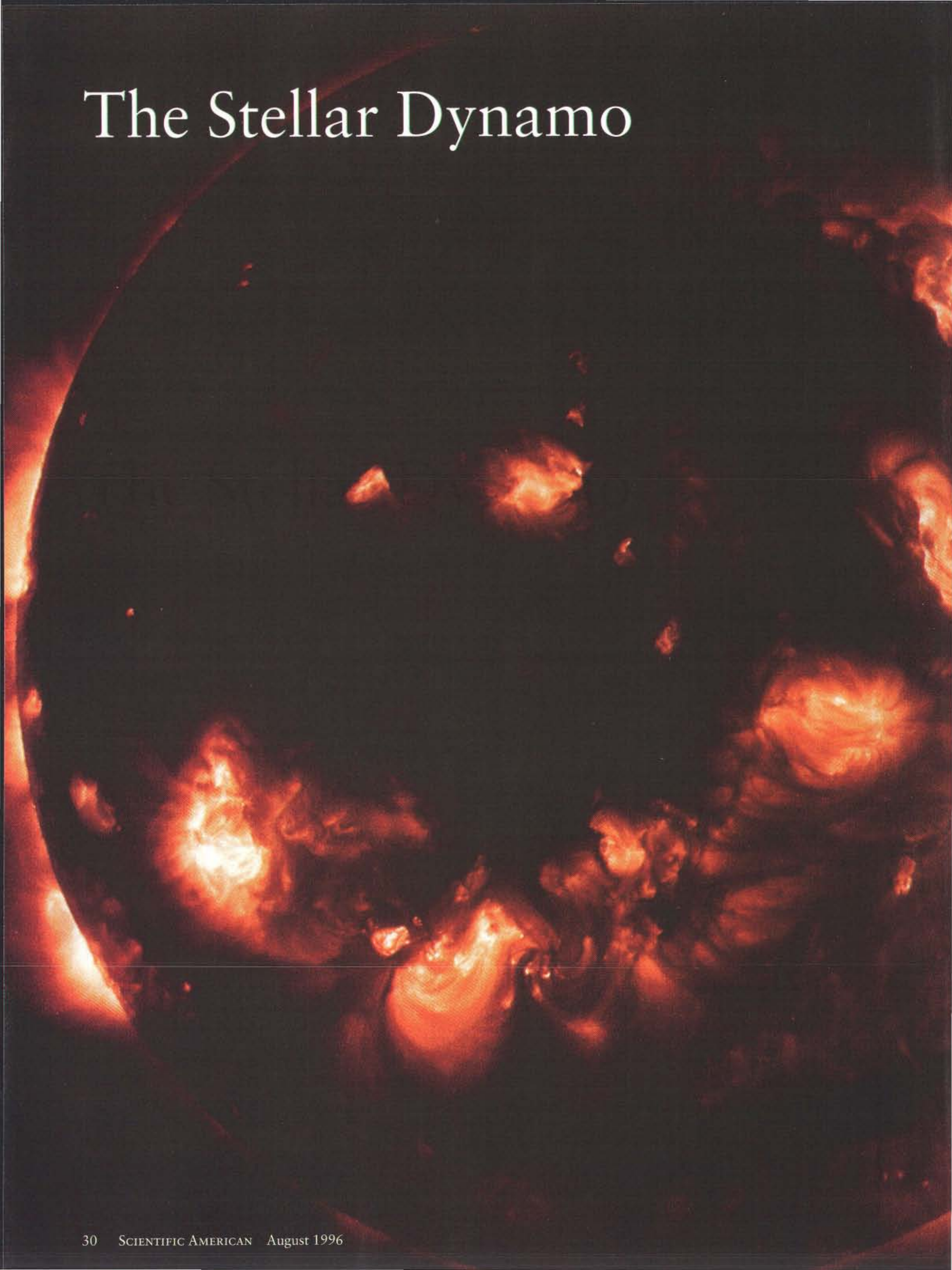
GET SET! SMARTCARDS ARE COMING TO AMERICA. Patrick Gauthier in *Portable Design,* Vol. 1, No. 6, pages 31–34; May 1996.

A CHIP OFF THE OLD SECURITY BLOCK. Andrea McKenna Findlay in *Card Technology* (Faulkner & Gray), Vol. 1, No. 2, pages 52–60; May/June 1996.

CRYPTOGRAPHIC SMART CARDS. David Naccache and David M'Raihi in *IEEE Micro,* Vol. 16, No. 3, pages 14–24; June 1996.

PUBLIC-KEY SECURITY SYSTEMS. Mahdi Abdelguerfi, Burton S. Kaliski, Jr., and Wayne Patterson in *IEEE Micro,* Vol. 16, No. 3, pages 10–13; June 1996.

The Stellar Dynamo



IBM RESEARCH AND SMITHSONIAN ASTROPHYSICAL OBSERVATORY

Sunspot cycles—on other stars—are helping astronomers study the sun's variations and the ways they might affect the earth

by Elizabeth Nesme-Ribes, Sallie L. Baliunas
and Dmitry Sokoloff

In 1801, musing on the vagaries of English weather, the astronomer William Herschel observed that the price of wheat correlated with the disappearance of sunspots. But the pattern soon vanished, joining what scientists at large took to be the mythology connecting earthly events with solar ones. That the sun's brightness might possibly vary, and thereby affect the earth's weather, remained speculative.

Thus, in the mid-1980s, when three solar satellites—*Solar Maximum Mission, Nimbus 7* and *Earth Radiation Budget*—reported that the sun's radiance was declining, astronomers assumed that all three instruments were failing. But the readings then perked up in unison, an occurrence that could not be attributed to chance. The sun was cooling off and heating up; furthermore, the variation was connected with the number of spots on its face.

In recent years one of us (Baliunas) has observed that other stars undergo rhythmic changes much like those of our sun. Such studies are helping refine our understanding of the "dynamo" that drives the sun and other stars. Moreover, they have revealed a strong link between "star spots" and luminosity, confirming the patterns discovered in our sun. But astrophysicists, including the three of us, are still debating the significance of the sun's cycles and the extent to which they might influence the earth's weather.

Sunspots

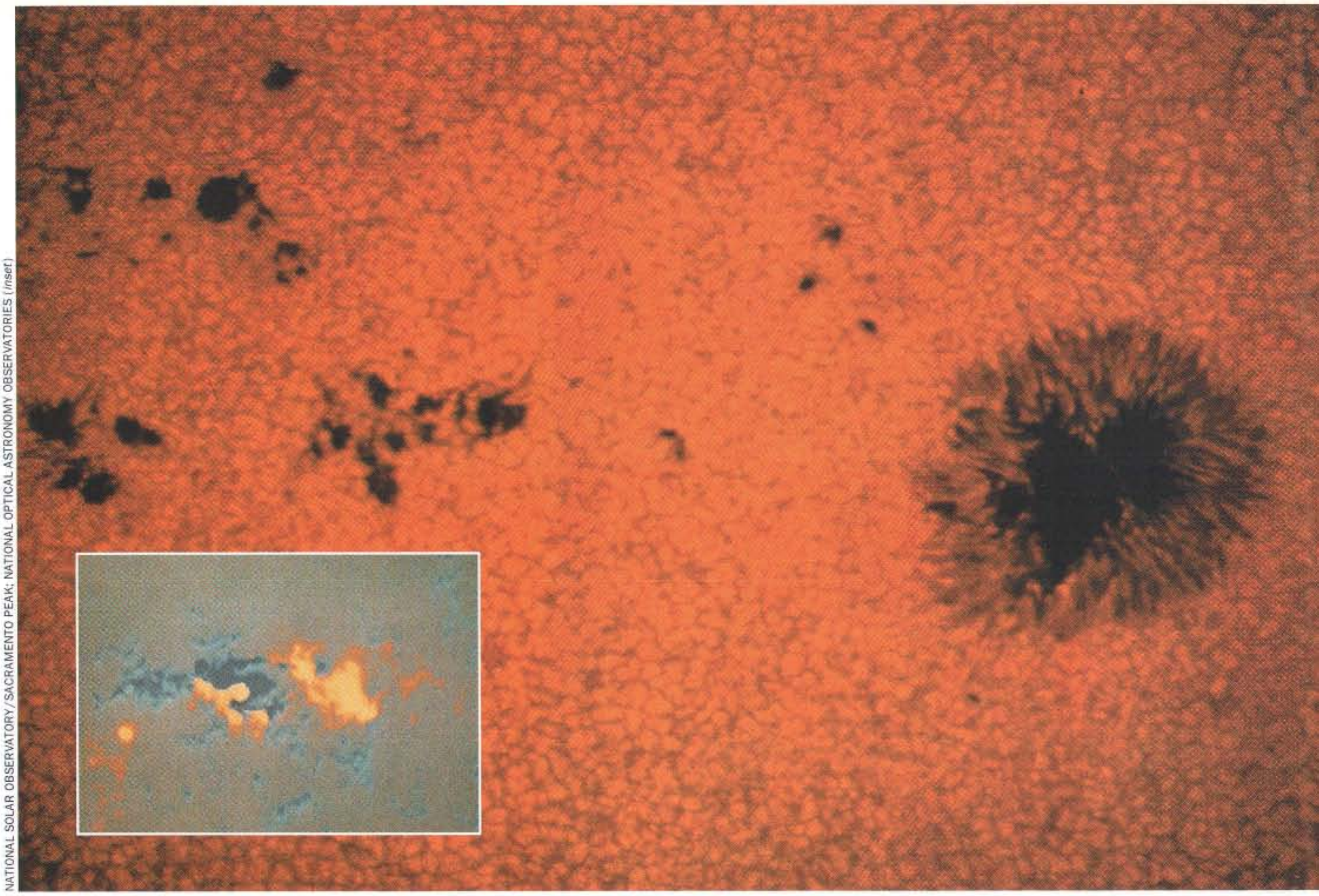
The earliest known sunspot records are Chinese documents that go back 2,000 years, preserving observations made by the naked eye. From 1609 to 1611 Johannes Fabricius, Thomas Harriot, Christoph Scheiner and Galileo Galilei, among others, began telescopic studies of sunspots. These records, as the German astronomer Samuel Heinrich Schwabe announced in 1843, displayed a prominent periodicity of roughly 10 years in the number of observed sunspot groups. By the 20th century George Ellery Hale of the Mount Wilson Observatory in California found those dark surface irregularities to be the seat of intense magnetic fields, with strengths of several thousand gauss. (The earth's magnetic field is, on the average, half a gauss.)

Sunspots appear dark because they are 2,000 degrees Celsius cooler than the surrounding surface of the sun; they would glow orange-red if seen against the night sky. The spots form when strong magnetic fields suppress the flow of the surrounding gases, preventing them from carrying internal heat to the surface. Next to the sunspots are often seen bright areas called plages (after the French word for "beach"). The magnetic-field lines tend to emerge from the surface at one spot to reenter the sun at another, linking the spots into pairs that resemble the two poles of a bar magnet that is oriented roughly east-west.

At the start of each 11-year cycle, sunspots first appear at around 40 degrees latitude in both hemispheres; they form closer to the equator as the cycle progresses. At sunspot minimum, patches of intense magnetism, called active regions, are seen near the equator. Aside from the sunspots, astronomers have observed that the geographic poles of the sun have weak overall magnetic fields of a few gauss. This large-scale field has a "dipole" configuration, resembling the field of a bar magnet. The leading sunspot in a pair—the one that first comes into view as the sun rotates from west to

MAGNETIC FIELDS on the sun are rendered visible in this x-ray photograph by the curving contours of solar flares. The lines of magnetic fields erupt from the sun's surface and heat the gases of the surrounding corona to up to 25 million degrees Celsius, causing them to glow. Flares are more frequent during sunspot maxima.

NATIONAL SOLAR OBSERVATORY/SACRAMENTO PEAK; NATIONAL OPTICAL ASTRONOMY OBSERVATORIES (*inset*)

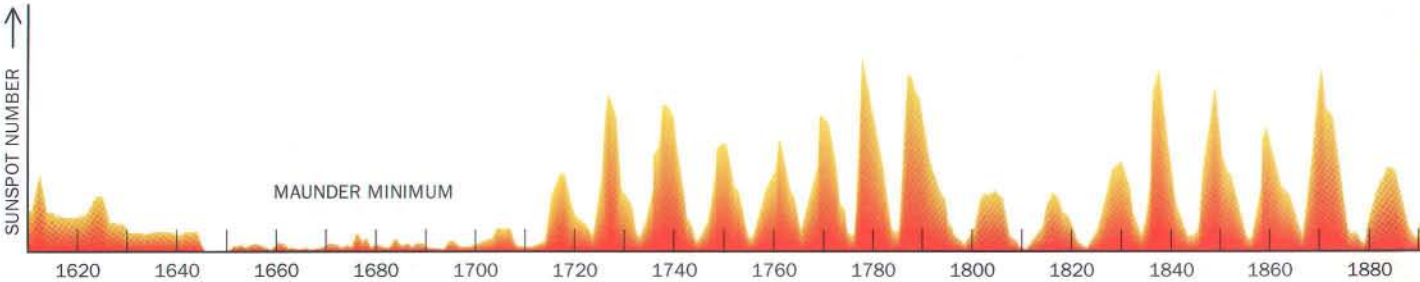


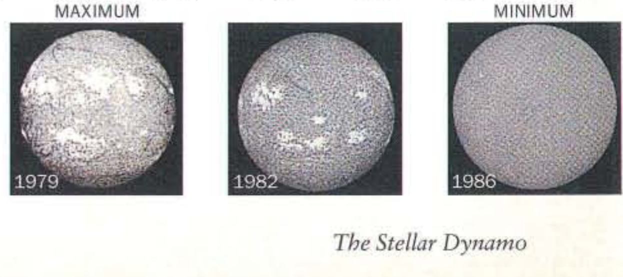
SUNSPOTS are relatively cool regions formed wherever magnetic fields emerge from the sun, thereby suppressing the upwelling of hot gases from the interior. Elsewhere on the surface, tightly coiled cells of cyclonically flowing gases show up as granules. Near a sunspot the magnetic fields organize the gaseous flow into lines resembling iron filings near a bar magnet. The magnetogram (*inset*) shows field lines emerging at one sunspot (*yellow*) and reentering at another (*blue*); such sunspot pairs are common.

east—has the same polarity as the pole of its hemisphere; the trailing sunspot has the opposite polarity. Moreover, as Hale and Seth B. Nicholson had discovered by 1925, the polarity patterns reverse every 11 years, so that the total magnetic cycle takes 22 years to complete.

But the sun's behavior has not always been so regular. In 1667, when the Paris Observatory was founded, astronomers there began systematic observations of the sun, logging more than 8,000 days of observation over the next 70 years. These records showed very little sunspot activity. This important finding did not raise much interest until the sunspot cycle was discovered, prompting Rudolf Wolf of Zürich Observatory to scruti-

NOAA; PETER SAMEK *Slim Films*





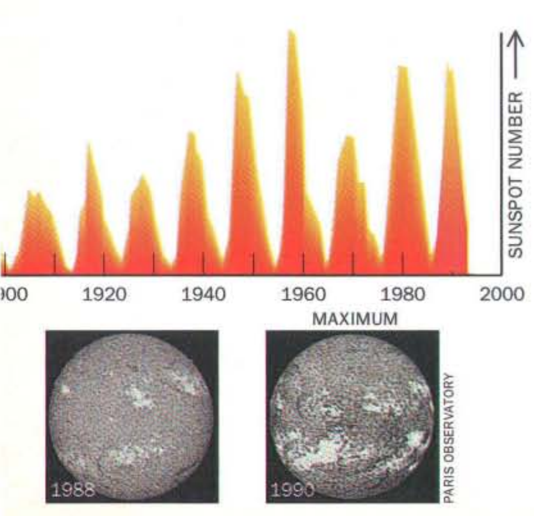
ELEVEN-YEAR CYCLES of sunspot activity were interrupted between 1645 and 1715 by a period of quiescence. This dearth of sunspots, called the Maunder minimum, coincided with unusually cool temperatures across northern Europe, indicating that solar fluctuations influence the earth's climate. The present regular pulsing of the sun's activity (*right*) was observed over one cycle at the Paris Observatory. These photographs were taken in violet light emitted by ionized calcium.

nize the records. Although he rediscovered the sunspot lull, Wolf's finding was criticized on the grounds that he did not use all the available documents.

During the late 1880s, first Gustav F.W. Spörer and then E. Walter Maunder reported that the 17th-century solar anomaly coincided with a cold spell in Europe. That astonishing observation lay neglected for almost a century, with many astronomers assuming that their predecessors had not been competent enough to count sunspots. It was only in 1976 that John A. Eddy of the University Corporation for Atmospheric Research in Boulder, Colo., reopened the debate by examining the Paris archives and establishing the validity of what came to be known as the Maunder minimum.

Eddy also noted that the amount of carbon 14 in tree rings increased during the dearth of sunspots. This radioactive element is created when galactic cosmic rays transmute nitrogen in the upper atmosphere. Eddy's findings suggested that when the magnetic fields in the solar wind—the blast of particles and energy that flows from the sun—are strong, they shield the earth from cosmic rays, so that less carbon 14 forms; the presence of excess carbon 14 indicated a low level of magnetic activity on the sun during the Maunder phase. Eddy thus reinforced the connection between the paucity of sunspots and a lull in solar activity.

Aside from the rarity of sunspots during the Maunder minimum, the Paris archives brought to light another oddity: from 1661 to 1705, the few sunspots that astronomers sighted were usually in the southern hemisphere. They were also traveling much more slowly across the sun's face than present-day sunspots do. Only at the beginning of the 18th century did the sun assume its modern appearance, having an abundance of sunspots rather evenly distributed between the two hemispheres.



The Solar Dynamo

The magnetic activity of the sun is believed to reside in its convective zone, the outer 200,000 kilometers where churning hot gases bring up energy from the interior. The fluid forms furious whorls of widely different sizes: the best known is an array of convective cells or granules, each 1,000 kilometers across at the surface but lasting only a few minutes. There are also "supergranules" that are 30,000 to 50,000 kilometers across and even larger flows. Rotation gives rise to Coriolis forces that make the whorls flow counterclockwise in the northern hemisphere (if one is looking down at the surface) and clockwise in the southern hemisphere; these directions are called cyclonic.

Whether similar cyclones exist underneath the surface is not known. Deep within, the convective zone gives way to the radiative zone, where the energy is transported by radiation. The core of the sun, where hydrogen fuses into helium to fuel all the sun's activity, seems to rotate rigidly and slowly compared with the surface.

The first description of how the sun's gases conspire to create a magnetic field was proposed in 1955 by Eugene N. Parker of the University of Chicago. Because of the high temperature, the atoms of hydrogen and helium lose their electrons, giving rise to an electrically charged substance, or plasma. As the charged particles move, they generate magnetic fields. Recall that the lines describing magnetic fields form continuous loops, having no beginning or end—their density (how closely together the lines are packed) indicates the intensity of the magnetic field, whereas their orientation reveals the direction. Because plasma conducts electricity very efficiently, it tends to trap the field lines: if the lines were to move through the plasma, they would generate a large, and energetically expensive, electric current.

Thus, the magnetic fields are carried along with the plasma and end up getting twisted. The entwined ropes wrap together fields of opposite polarity, which tend to cancel each other. But the sun's rotation generates organizational forces that periodically sort out the tangles and create an overall magnetic field. This automatic engine, which generates magnetism from the flow of electricity, is the solar dynamo.

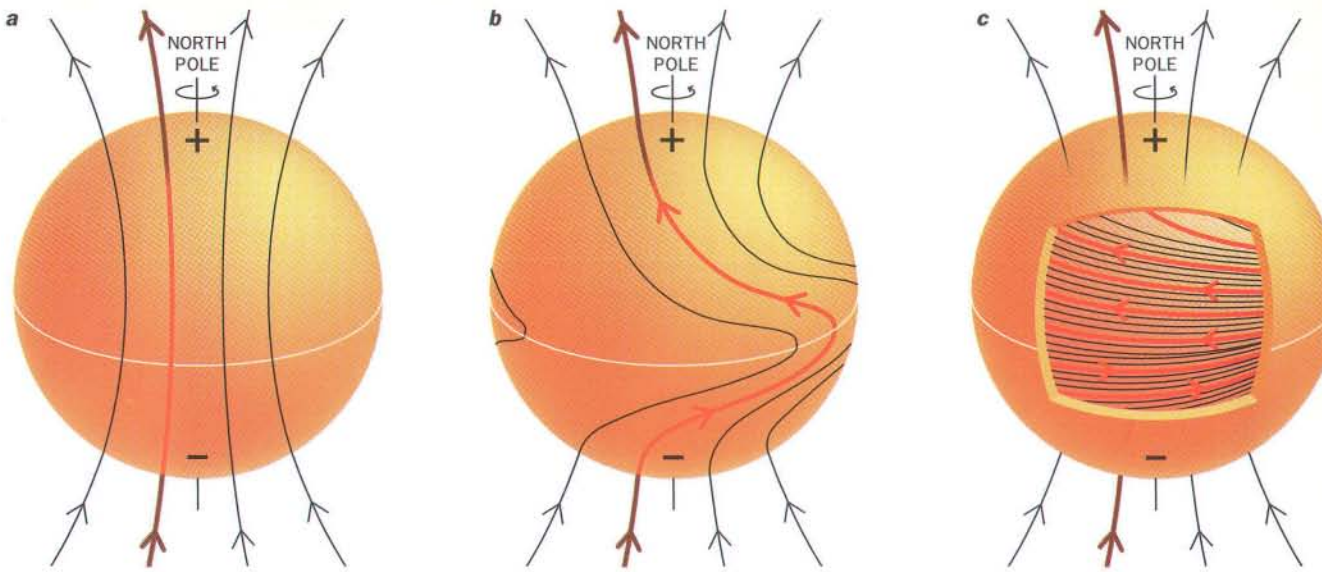
The dynamo has two essential ingredients: the convective cyclones and the sun's nonuniform rotation. During the mid-1800s, Richard C. Carrington, an English amateur astronomer, found that the sunspots near the equator rotate faster, by 2 percent, than those at midlatitudes. Because the spots are floating with the plasma, the finding indicates that the sun's surface rotates at varying speeds. The rotation period is roughly 25 days at the equator, 28 days at a latitude of 45 degrees and still longer at higher latitudes. This differential rotation should extend all the way through the convective zone.

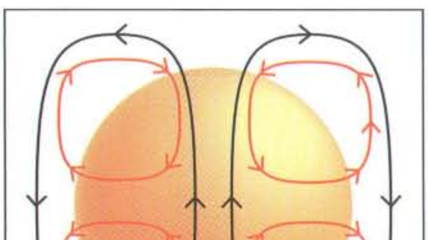
Now suppose that the initial shape of the sun's field is that of a dipole oriented roughly north-south. The field lines get pulled forward at the equator by the faster rotation and are deformed in the east-west direction. Ultimately, they lie parallel to the equator and float to the surface, erupting as pairs of sunspots.

But Coriolis forces tend to align the cyclones and thereby the sunspots, which are constrained to follow the plasma's gyrations. The cyclones arrange the sunspots so that, for example, a trailing sunspot in the northern hemisphere lies at a slightly higher latitude than a leading one. As the equatorial field lines are stretched, they eventually unwind and drift outward. The trailing sunspot reaches the pole first, effectively reversing the magnetic field there. (Recall that the trailing spot has a polarity opposite that of the nearest pole.) Those field lines that initially extended far beyond the sun reconnect into loops and are blown away by the solar wind. In this manner, the overall magnetic field flips, and the cycle begins again.

There is, however, a caveat. This simple picture seems to be at odds with recent results from helioseismology, the science of sunquakes. The model requires the sun to rotate faster at the interior; in contrast, results from the Global Oscillation Network Group, an international collaboration of observatories, show that the rotation velocity near the equator decreases downward. Such experiments are providing accurate information on internal motions of the sun and thereby helping to refine dynamo theory.

But what happened during the Maunder minimum? To explain this lull, two of us (Nesme-Ribes and Sokoloff) not-





SOLAR DYNAMO generates the sun's magnetic field and also causes it to change orientation every 11 years. Suppose that the initial magnetic field (*a*) resembles that of a bar magnet with its north pole (+) near the sun's geographic north pole. The magnetic-field lines are carried along with the electrically charged gases. The faster flow at the equator therefore distorts the field lines (*b*) until they wrap tightly (*c*) around the sun.

ed that apart from a dipole pattern, the magnetic field must also have a small quadrupole component, resembling the field of two bar magnets placed side by side. If the quadrupole oscillates at a slightly different rate than the dipole, the sunspots in one hemisphere are produced slightly earlier than those in the other hemisphere—precisely what we observe now. Furthermore, over the past four centuries, a few solar cycles showed different numbers of sunspots in the northern and southern hemispheres. This pattern seems to repeat every century or so, exactly what one would expect if the dipole "beats" with a weak quadrupole.

But suppose that the quadrupole field is as strong as the dipole. The equatorial field lines that result from stretching this combination will then cancel out in one hemisphere yet remain in the other. The few spots that do appear will all be in one hemisphere, just as 17th-century astronomers noted during the Maunder minimum.

We can encapsulate the complex relation between the dipole and quadrupole fields in a "dynamo number" D. This number is the product of the helicity, or spiraling motion, of the plasma and the local rate of change of rotation. When D is very small, the magnetic field tends to die out; as it increases, however, the quadrupole field shows up, with the dipole following. Beyond a critical value of D, both components of the field are steady. But as D increases further, the dynamo becomes periodic, increasing and decreasing in regular cycles; this is the regime in which the sun now lies. A weak quadrupole field, beating in phase with the dipole, leads to short and intense cycles; a stronger quadrupole field, if slightly out of phase with the dipole field, lengthens and weakens the sunspot cycle. Far beyond the critical dynamo number, chaos results.

Dynamic Stars

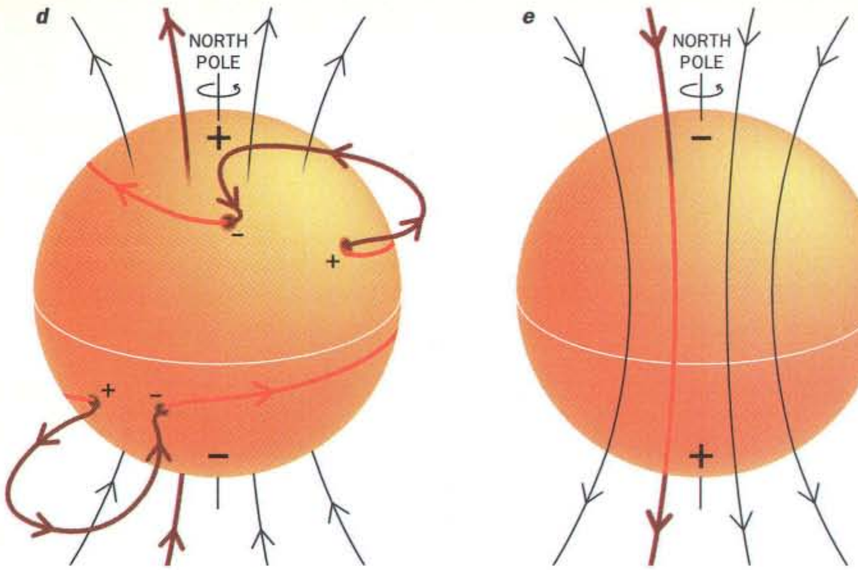
As we now know, the sun's brightness increases with the magnetic activity over a cycle: the bright plages overwhelm the dark sunspots. (Presumably, as the sun brightens and darkens, its total energy is temporarily channeled into different reservoirs—kinetic, magnetic, thermal or potential.) During the past 16 years of satellite observations, the sun's total energy output has varied roughly 0.1 percent between a brighter, magnetically active phase and a fainter, quiet one.

Because of the brevity of the satellite records, we do not know the variability of the sun's brightness over decades. This value, however, is vital to evaluating the sun's influence on the earth. One possible way to answer this question is to examine "star spot" cycles on other stars.

It is not easy to map the features on the surface of stars. But as magnetic fields heat the outer layers of a star's atmosphere, they radiate the energy in certain spectral lines. For example, on our sun, the intensity of the two violet emission lines of calcium (having wavelengths of 396.7 and 393.4 nanometers) closely follows the strength and extent of the magnetic fields. Variations in these lines thus give us a measure of the changing surface magnetism of a star.

At Mount Wilson Observatory in 1966, Olin C. Wilson began a program of measuring the magnetic activity of roughly 100 so-called main-sequence stars—those that, like the sun, are burning hydrogen. (When the hydrogen runs out, a star expands into a red giant.) Most of these stars show obvious signs of magnetic activity, by way of variations in their violet calcium emission lines. The fluctuations vary greatly in amplitude and duration, depending primarily on the age and mass of the star.

All these stars have a dynamo number higher than the critical value required for sustaining magnetic fields. For a young star of one or two billion years, the rotation period is fast, roughly 10 to 15 days. The resulting high value of



PETER SAMEK *Slim Films*

But the field lines then resist the stretching and unwind, moving up toward the surface and erupting as sunspot pairs (*d*). The sunspots drift toward the poles, with the trailing sunspot reaching first; as a result, the overall field flips (*e*). In addition to the dipole field above, the sun probably also has a "quadrupole" field (*opposite page, red*) whose "beating" with the dipole field was responsible for the Maunder minimum.

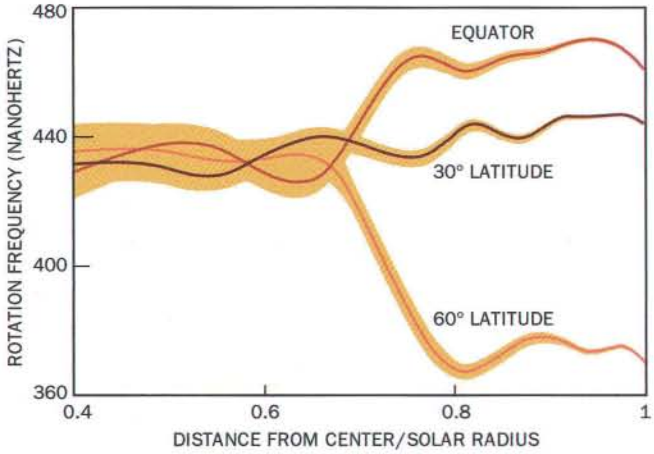
D means that these young stars have erratic fluctuations in magnetic activity over intervals as short as two years and no well-defined cycles. The fluctuations sometimes repeat, however, having periods between two and 20 years or so that lengthen with age.

But as a star ages, it slows down—because its angular momentum is carried off by the magnetic wind—and *D* falls. Then a consistent dynamo cycle begins to appear, with a period of about six to seven years and sometimes even with two independent periods. Later on—for even lower *D*—one period starts to dominate, lengthening with age from eight to 14 years. In addition, there are occasional Maunder minima. If rotation were to slow further, in the very oldest stars, we predict that the magnetic field should be steady. The Wilson sample contains a few very old stars, but they still show cycles, indicating that the steady dynamo would not be reached in 10 billion years—soon after which they will expand into red giants.

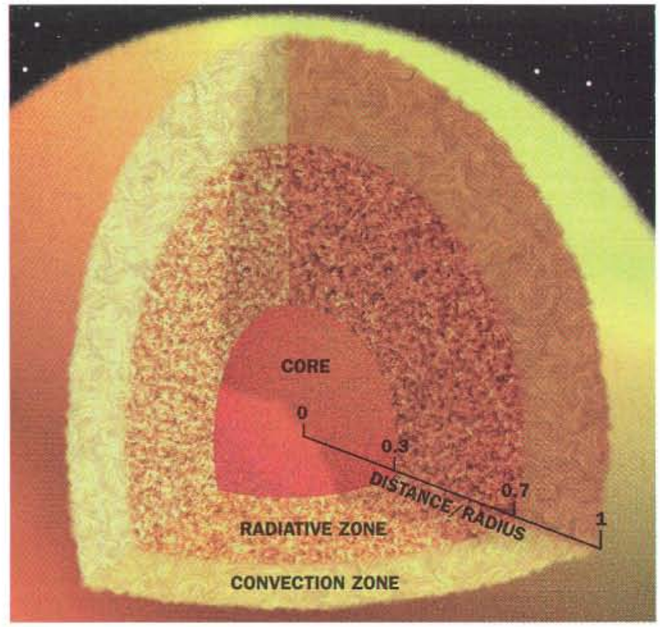
To focus on the solar dynamo, we (Baliunas and her collaborators at Mount Wilson and Tennessee State University) restricted Wilson's broad sample of stars to those similar to our sun in mass and age. That group currently comprises 30-year records of 20 to 30 stars, depending on the criteria defining similarity to the sun. Most of these stars show prominent cycles similar in amplitude and period to those of the sun. About one fourth of the records show that the stars are in a dead calm, suggesting a phase similar to our sun's Maunder minimum. This finding implies that sunlike stars spend a fourth of their lives in a lull.

We have just discovered one star, HD 3651, in transition between the cyclic and Maunder minimum phase. HD 3651 showed periodic behavior for about 12 years and then stopped fluctuating as its surface activity dropped to very low levels. Its entry into the Maunder minimum state was surprisingly rapid. Thus do sunlike stars, observed over a few decades, offer us "snapshots" of the range of solar variability over timescales of centuries.

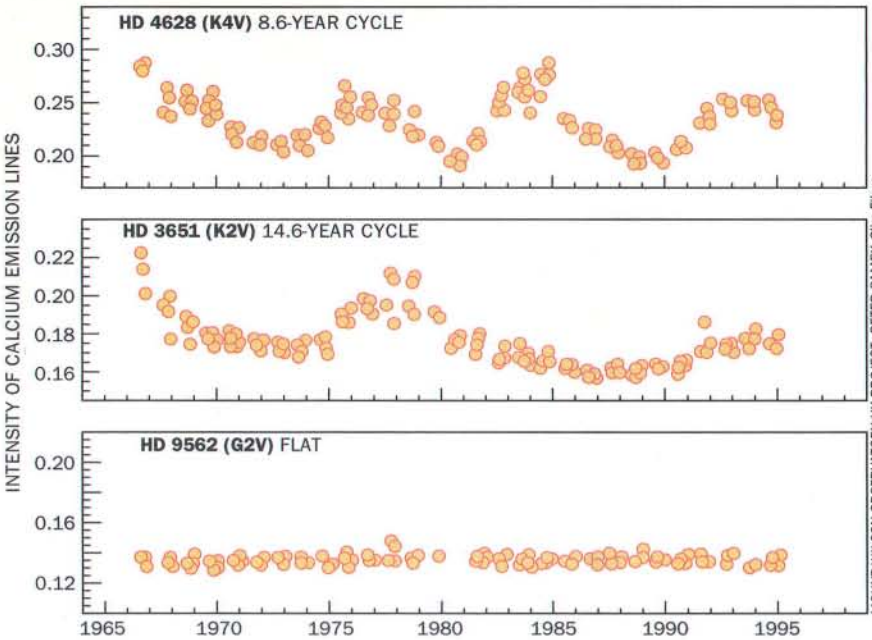
The brightness of these sunlike stars can also be compared with their magnetic activity. In 1984 thorough and precise photometric observations of some of the Wilson stars began at the Lowell and Sacramento Peak observatories. Since 1992 those of us at Tennessee State and the Smithsonian Astrophysical Observatory have used automated telescopes to observe some of these stars. All the stars are brightest near the peak of the activity cycle. Some stars vary as little as our sun does—only 0.1 percent over the last 11-year cycle—but other



PETER SAMEK *Slim Films*



ROTATION of the sun's surface is faster at the equator and slower near the poles. This differential rotation (as measured by means of sunquakes by the Global Oscillation Network Group) extends through the outer layers. The sun's core, in which fusion generates the energy that ultimately powers the dynamo, most likely rotates at a constant angular velocity, like a rigid body.



STAR CYCLES—changes in magnetic activity in nearby stars—are detected via their violet calcium emission lines. These three stars show magnetic behavior resembling that of our sun: steady cycles (*top*), cycles subsiding into a Maunder-type minimum (*middle*) and dead quiet, implying a Maunder phase (*bottom*). The study of these stars indicates that the sun is capable of far greater variation than it currently exhibits.

sunlike stars have varied by as much as 0.6 percent in a cycle. Thus, the sun's current behavior might be a poor indicator of the full range of fluctuations of which it is capable.

Over the decades, researchers have inferred the evolutionary history of a sunlike star from the collection of stellar records. A young star has a relatively rapid rotation period of several days and high, irregular levels of surface magnetism. Changes in brightness of several percent accompany the magnetic variations. The young star is, however, darkest during the peak of magnetic activity, presumably because the dark spots are so large that they, not the plages, dominate. As the sunlike star ages, it rotates more slowly, and the magnetic activity decreases. Maunder minima appear in these "older" stars; furthermore, radiance now peaks at sunspot maximum, with fluctuations of 1 percent or less over a cycle.

Influencing the Earth

The star-spot results point to a change in brightness of at least 0.4 percent between the cyclic phase and the Maunder minimum phase. This value corresponds to a decrease in the sun's net energy input of one watt per square meter at the top of the earth's atmosphere. Simulations performed at the Laboratory of Dynamic Meterology in Paris and elsewhere suggest that such a reduction, occurring over several decades, is capable of cooling the earth's average temperature by 1 to 2 degrees C—enough to explain the observed cooling during the Maunder minimum.

But greenhouse gases made by humans may be warming the earth, by trapping heat that would otherwise radiate into space. This warming is equivalent to the earth's receiving radiation of two watts per square meter at the surface. The sun has apparently delivered to the earth no more or less than 0.5 to 1.0 watt per square meter over the past few centuries. Therefore, if direct heating is the only way in which the sun affects the earth's climate, the greenhouse gases should already be dominating the climate, washing out any correlations with the sun's activity.

The link between climate and sunspots seems, however, rather persistent. The length of the sunspot cycle, for example, is closely correlated with global temperatures of the past 100 years. Six out of seven minima in solar magnetism during the past 5,000 to 6,000 years, as traced by radiocarbon in tree rings, coincide with intervals of cooler climate. In addition, the sunspot cycle correlates with stratospheric wind patterns, for reasons not yet well understood. All these pieces of evidence induce some scientists, including us, to argue that the sun must also be influencing the earth by powerful indirect routes.

Variations in the sun's ultraviolet radiation, for example, may be changing the ozone content of our upper atmosphere, as well as its dynamics. Recent simulations also indicate that winds in the lower stratosphere can convey variations in solar radiance down to the troposphere, where they interact more directly with the weather system. Such matters are now the subject of vigorous debate. Unraveling the ways in which the sun warms the earth provides vital information concerning the role played by humankind—and the role played by the sun—in the process of climatic change. SA

The Authors

ELIZABETH NESME-RIBES, SALLIE L. BALIUNAS AND DMITRY SOKOLOFF all are active in unraveling connections between the sun's variations and the earth's climate. Nesme-Ribes is an astronomer at the Paris Observatory and the National Center for Scientific Research in France. Apart from studying the solar dynamo, she has conducted extensive searches into the 17th-century archives on sunspots at her home institution. Baliunas is a scientist at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Mass. She observes the variations of sunlike stars at the Mount Wilson Observatory in Pasadena, Calif., where she is deputy director. Sokoloff is professor of mathematics in the department of physics at Moscow State University in Russia.

Further Reading

THE VARIABLE SUN. Peter V. Foukal in *Scientific American*, Vol. 262, No. 2, pages 26–33; February 1990.

GUIDE TO THE SUN. Kenneth J. H. Phillips. Cambridge University Press, 1992.

EXPLORER OF THE UNIVERSE: A BIOGRAPHY OF GEORGE ELLERY HALE. Helen Wright. American Institute of Physics Press, 1994.

SUN, EARTH AND SKY. Kenneth R. Lang. Springer-Verlag, 1995.

Gradients That Organize Embryo Development

A few crucial molecular signals give rise to chemical gradients that organize the developing embryo

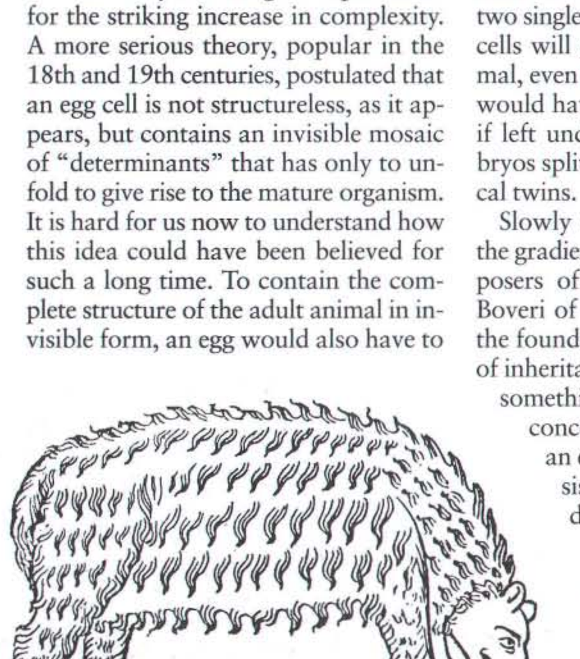
by Christiane Nüsslein-Volhard

Bears mate in wintertime. The female then retires into a cave to give birth, after several months, to three or four youngsters. At the time of birth, these are shapeless balls of flesh, only the claws are developed. The mother licks them into shape.

This ancient theory, recounted by Pliny the Elder, is one of the many bizarre early attempts to explain one of life's greatest mysteries—how a nearly uniform egg cell develops into an animal with dozens of types of cells, each in its proper place. The difficulty is finding an explanation for the striking increase in complexity. A more serious theory, popular in the 18th and 19th centuries, postulated that an egg cell is not structureless, as it appears, but contains an invisible mosaic of "determinants" that has only to unfold to give rise to the mature organism. It is hard for us now to understand how this idea could have been believed for such a long time. To contain the complete structure of the adult animal in invisible form, an egg would also have to contain the structures of all successive generations, because adult females will in time produce their own eggs, and so on, ad infinitum. Even Goethe, the great poet and naturalist, favored this "preformation hypothesis," because he could not think of any other explanation.

About 100 years ago experimental embryologists began to realize that developmental pathways need not be completely determined by the time the egg is formed. They discovered that some experimental manipulations led to dramatic changes in development that could not be explained by the mosaic hypothesis. If an experimenter splits a sea-urchin embryo at the two-cell stage into two single cells, for example, each of the cells will develop into a complete animal, even though together the two cells would have produced only one animal if left undisturbed. When human embryos split naturally, the result is identical twins.

Slowly an important idea emerged: the gradient hypothesis. One of the proposers of this idea was Theodor H. Boveri of the University of Würzburg, the founder of the chromosomal theory of inheritance. Boveri suggested that "a something increases or decreases in concentration" from one end of an egg to the other. The hypothesis, in essence, is that cells in a developing field respond to a special substance—a morphogen—the concentration of which gradually increases in a certain direction, forming a gradient. Different concentrations of the morphogen were postulated to cause different responses in cells.



COURTESY OF CHRISTIANE NÜSSLEIN-VOLHARD

SEEMING MIRACLE of animal development confounded early scholars. This 16th-century drawing shows a bear licking into shape her offspring, which were believed to be born shapeless.

Although concentration gradients of morphogens could in principle explain how cells "know" their position in an embryo, the idea was for a long time not widely accepted. One of the difficulties lay in explaining how a morphogenetic gradient could be established and then remain stable over a sufficient period. In a developing tissue composed of many cells, cell membranes would prevent the spread of large molecules that might form a concentration gradient. In a single large egg cell, conversely, diffusion would quickly level such a gradient. Further, the biochemical nature and the mechanism of action of morphogens were a mystery.

For most biologists, the means of gradient formation remained elusive until recently, when researchers in several laboratories discovered gradients operating in the early embryo of the fruit fly, *Drosophila*. For most nonbiologists, it is a surprise that many of the mechanisms of development are best known in *Drosophila*, rather than in animals more closely related to humans. The examples I shall describe illustrate the reason for the preeminence of *Drosophila* as an experimental subject: a lucky coincidence of advantages makes it almost ideal for studies in genetics, embryology and molecular biology.

Drosophila became the laboratory animal of choice for studying Mendelian genetics early this century because the fly is easy to handle and quick to breed

JEREMY BURGESS *Science Photo Library, Photo Researchers, Inc.*

in large numbers, making it possible to search through many individual flies for mutants. Studies of mutants have successfully elucidated metabolic pathways and regulatory processes in viruses, bacteria and yeast. Twenty years ago Eric F. Wieschaus, now at Princeton University, and I extended this approach to *Drosophila* by searching for genes that control the segmented form of the larva. The larva is relatively large—about one millimeter long—and has well-defined, repeated segments that emerge within 24 hours of the laying of the egg. These features are crucial for interpreting experimentally induced abnormalities that affect the pattern of development.

Another key advantage of using *Drosophila* for embryological studies is that during its early development the embryo is not partitioned into separate cells. In the embryos of most animals, when a cell's nucleus divides, the rest of the cell contents divides with it. Cell membranes then segregate the halves, yielding two cells where there was one. Hence, the embryo grows as a ball of cells. In contrast, the nucleus of the fertilized *Drosophila* egg divides repeatedly, but membranes do not isolate the copies. Eventually thousands of nuclei lie around the periphery of what is still, in a manner of speaking, a single cell. Only after three hours of cell division, when some 6,000 nuclei have formed, do separating membranes appear.

This peculiarity allows chemicals to spread freely through the early embryo and influence the developmental fate of large regions of it. As experimentalists, we can therefore transplant cytoplasm (the viscous fluid within cells) or inject biological molecules into various regions of a *Drosophila* embryo and observe the results.

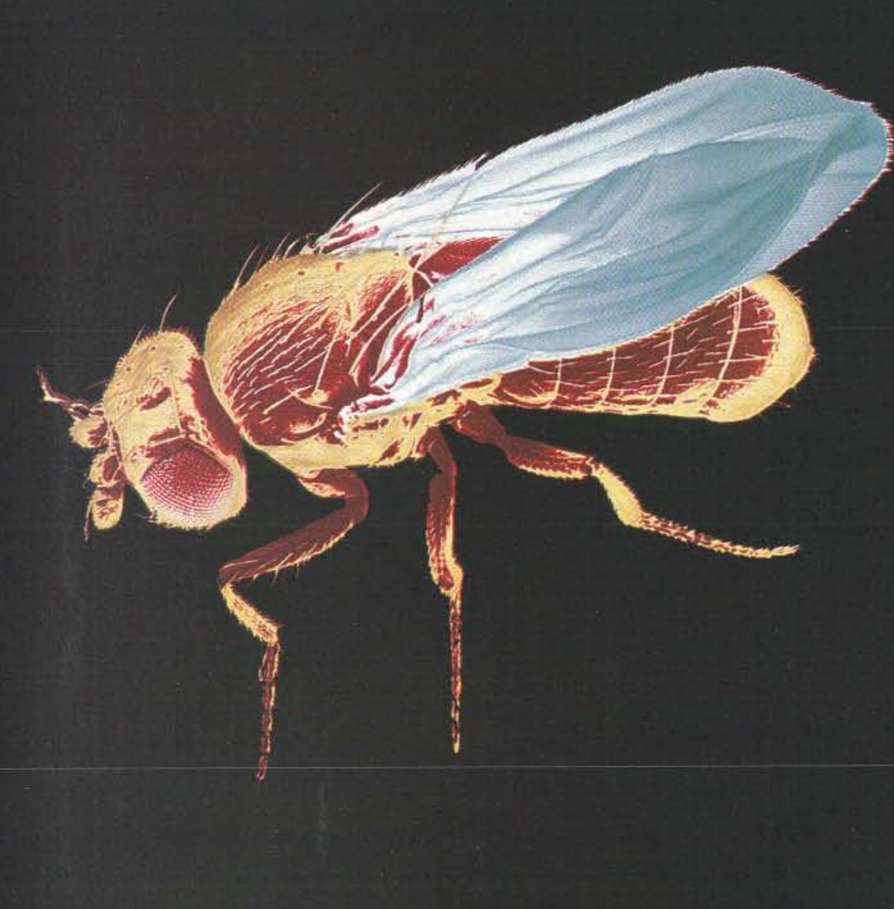
The Power of Gradients

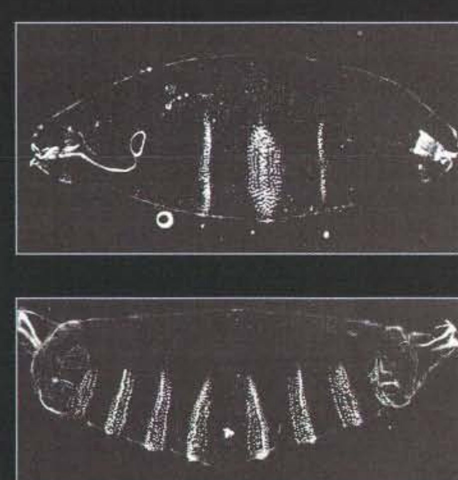
In addition, *Drosophila* is fairly easy to study with the techniques of molecular biology. The insect has only four pairs of chromosomes, and they exist in a special giant form. The giant chromosomes make it possible to see under the microscope, in many cases, the disruptions in the genetic material caused by mutations. This fact helps a great deal when the mutations are being studied. Last but not least, by exploiting naturally occurring mobile genetic elements, it is possible to add, with high efficiency, specific genes to the genetic complement of *Drosophila*.

By studying mutants, researchers have found about 30 genes that are active in the female and organize the pattern of the embryo. Only three of them encode molecular signals that specify the structures along the long antero-posterior (head-tail) axis of the larva. Each signal is located at a particular site in the developing egg and initiates the creation of a different type of morphogenetic gradient. In each case, the morphogen has its maximum concentration at the site of the signal.

One of the signals controls the development of the front half of the egg, which gives rise to the head and thorax of the larva. A second signal controls the region that develops into the abdomen, and the third controls development of structures at both extreme ends of the larva.

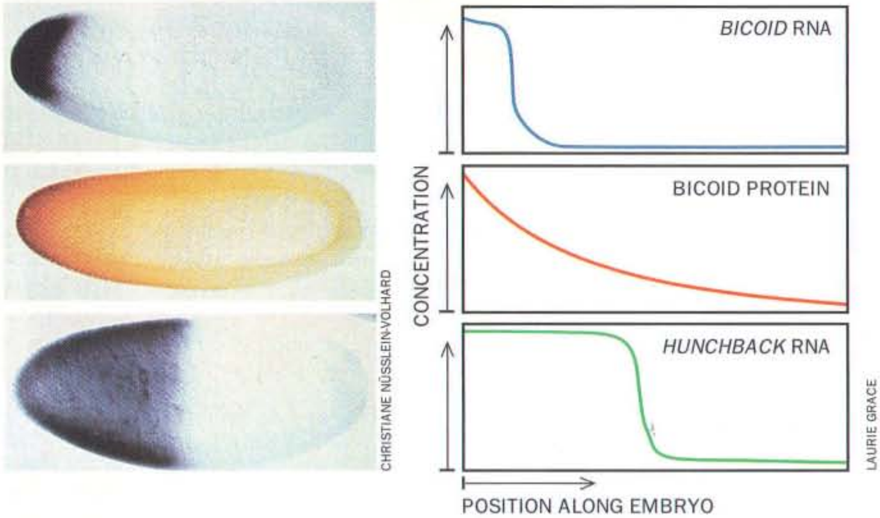
The simplest of the morphogenetic gradients initiated by these signals consists of a protein called Bicoid, which





CHRISTIANE NÜSSLEIN-VOLHARD

MANIPULATION of protein gradients has produced two abnormal embryos of the fruit fly *Drosophila melanogaster* (*left*). One has two head ends in mirror symmetry (*top*); the other has two abdominal ends (*bottom*). The embryos, which do not develop into viable larvae, are stained to show specific proteins.



FRESHLY LAID EGG of *Drosophila* has *bicoid* RNA localized at the anterior, or head, end (*visualized by staining at top left*). Two hours later Bicoid protein from this signal has spread along the embryo (*middle panels*). The Bicoid concentration gradient exceeds a threshold level and activates the *hunchback* gene only in the front half of the embryo (*bottom panels*).

determines the pattern in the front part of the larva. My colleague Wolfgang Driever and I found that a concentration gradient of Bicoid is present in the *Drosophila* embryo from the very earliest stages. The concentration is highest at the head end of the embryo, and it declines gradually along the embryo's length. Mutations in the *bicoid* gene of a *Drosophila* female prevent the development of a Bicoid gradient. The resulting embryos lack a head and thorax.

Bicoid acts in the nuclei of the embryo. The protein is termed a transcription factor, because it can initiate transcription of a gene. This is the process whereby messenger RNA (mRNA) is produced from the genetic material, DNA; the cell then uses the mRNA to synthesize the gene's protein product. Transcription factors operate by binding to specific DNA sequences in the control regions, or promoters, of target genes. In order to bind to a promoter, Bicoid must be present above a certain threshold concentration.

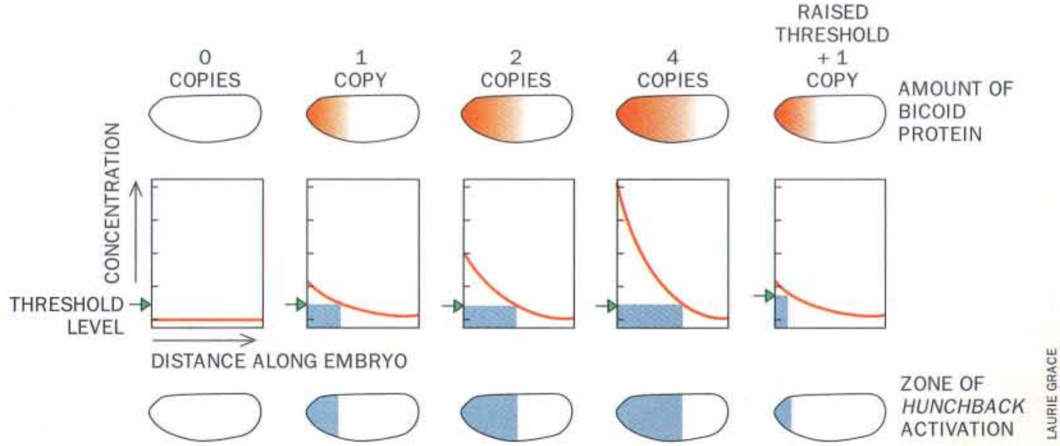
Driever and I have investigated the interaction of Bicoid with one target gene in particular, *hunchback*. *Hunchback* is transcribed in the front half of the early embryo, and the gene's promoter contains several Bicoid binding sites. We carried out two types of experiment: in one, we changed the concentration profile of Bicoid, and in the other we changed the structure of the *hunchback* gene promoter.

By introducing additional copies of the *bicoid* gene into the female, it is possible to obtain eggs with levels of Bicoid that are four times higher than normal all along the gradient. In these embryos, the zone of *hunchback* gene activation extends toward the posterior, and the head and thorax develop from a larger part of the embryo than is normal. This abnormality could in principle arise either because the Bicoid concentration gradient was steeper in the manipulated embryos or because the absolute level of Bicoid concentration was higher. The correct interpretation was made clear by an experiment in which we made mutant embryos that had a level Bicoid concentration along their length, so there was no gradient at all. These embryos produced only one type of anterior structure (head or thorax); which type depended on the Bicoid concentration. The experiment shows, then, that the absolute concentration of Bicoid, not the steepness of the gradient, is important for controlling subsequent development of each region.

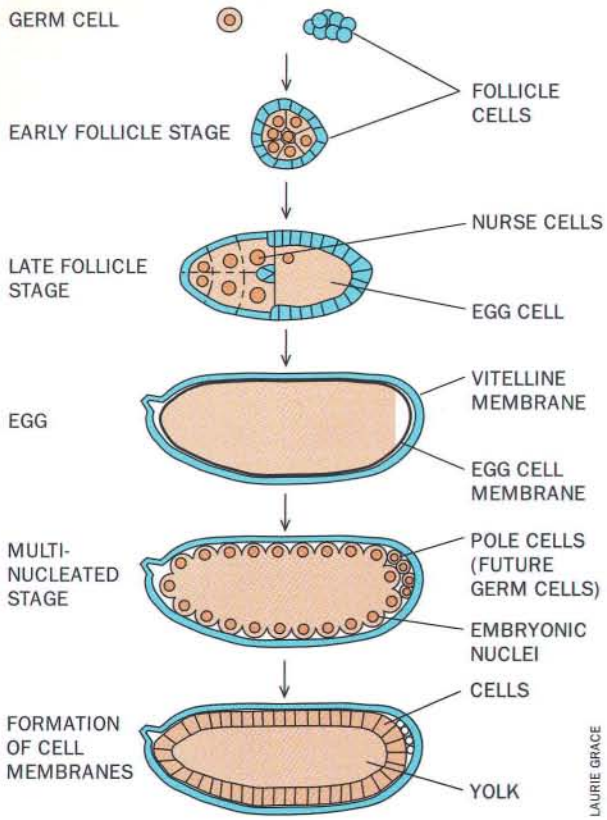
In the second type of experiment the Bicoid gradient was left unchanged, but the promoter region of the *hunchback* gene was altered. When the altered promoter bound only weakly to Bicoid, higher Bicoid concentrations were required to initiate *hunchback* transcription. Consequently, the edge of the zone of *hunchback* activity shifted forward. In these embryos, as one might predict, the head forms from a smaller than normal region. This experiment revealed that Bicoid exerts its effect by binding to the *hunchback* promoter.

These experiments show how a morphogen such as Bicoid can specify the position of a gene's activation in an embryo through its affinity for the gene, in this case *hunchback*. In theory, a large number of target genes could respond to various thresholds within the gradient of a single morphogen, producing many different zones of gene activation. In reality, however, a gradient acts directly on usually no more than two or three genes, so it specifies only two or three zones of activation.

How is the morphogenetic Bicoid gradient itself established? While the unfertilized egg is developing, special nurse cells connected to it deposit mRNA for



EMBRYOS with extra copies of the *bicoid* gene produce steeper gradients of Bicoid protein. The region where Bicoid concentration exceeds the threshold for activation of the *hunchback* gene then expands. If the activation threshold is artificially increased, the zone of *hunchback* activity shrinks.



***DROSOPHILA* EGG is built from a germ cell, with nurse cells and follicle cells providing nutrients and other factors that control embryonic development. Only after three hours, when there are 6,000 nuclei, do cell membranes form. Larval tissues appear later.**

Bicoid at its anterior tip. Synthesis of Bicoid, which starts at fertilization, is therefore already under way when the egg is laid. As the embryo develops, the protein diffuses away from the site of its production at the head end. Bicoid is unstable, however, so its concentrations at remote points—that is, at the end that will become the abdomen—never reach high levels. The resulting concentration gradient persists until cell membranes form.

This simple diffusion mechanism is accurate enough to meet the requirements of normal development. Remarkably, even substantial changes in Bicoid levels—doubling or halving—result in normally proportioned larvae. It appears that mechanisms operating at later stages of development can correct some errors in the early stages. If a researcher transplants *bicoid* mRNA into the posterior pole of a normal embryo, an additional Bicoid protein gradient arises, oriented opposite to the natural one. The resulting embryo displays a duplicate head where the abdomen should be. This experiment shows conclusively that *bicoid* mRNA is by itself sufficient to determine polarity.

Other work has revealed how the *bicoid* mRNA is positioned precisely within the egg cell. Paul M. Macdonald of Stanford University has identified a large section of the *bicoid* mRNA molecule that contains all the information required for a cell to recognize it, transport it and anchor it. Daniel St. Johnston and Dominique Ferrandon, while working in my laboratory, found that a molecular complex consisting of *bicoid* mRNA and a protein known as Staufen will move in one direction along structural elements in cells called microtubules. It seems likely that this effect explains the localization of *bicoid* mRNA, although other proteins are certainly also involved.

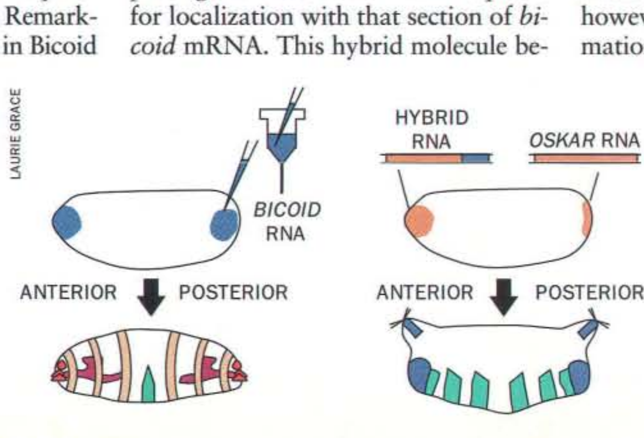
Whereas Bicoid is determining the anterior section of the larva's long axis, a different morphogenetic gradient is determining the posterior part. The gradient in this case is composed of the protein Nanos. *Nanos* mRNA localizes in the cytoplasm at the posterior end of the egg. This occurrence depends critically on another molecular complex consisting of the Staufen protein and mRNA from a gene named *oskar.* Anne Ephrussi and Ruth Lehmann, then at the Whitehead Institute for Biomedical Research in Cambridge, Mass., demonstrated the crucial role of *oskar* by replacing the section of mRNA required for localization with that section of *bicoid* mRNA. This hybrid molecule behaved like *bicoid* mRNA, collecting at the anterior pole rather than at the posterior one. The manipulation misdirected the *nanos* mRNA to the anterior pole, causing the embryos to develop with two abdominal ends in mirror symmetry.

Getting around Cell Membranes

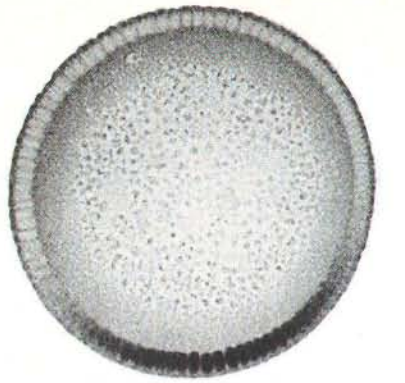
The mechanisms that produce the morphogenetic gradients of Bicoid and Nanos, both of which are large molecules, can operate only when there are no cell membranes to hinder diffusion. In most animals, however, early development creates cell membranes between different regions of the egg, so these mechanisms cannot work. It is notable, then, that the dorsoventral (top-bottom) axis of the *Drosophila* embryo, unlike the antero-posterior axis, is defined by a single gradient that could develop even in the presence of cell membranes. This mechanism may thus be more typical of those found in other creatures.

The first embryonic pattern along the dorsoventral axis is determined by the gradient of a protein called Dorsal. Like Bicoid, Dorsal is a transcription factor, and it controls the activity of several target genes in a concentration-dependent manner. The Dorsal protein acts as both a transcriptional activator and a repressor—inside cell nuclei, it turns genes on or off. When its concentration in the cell nucleus exceeds a particular threshold, Dorsal activates the transcription of a pair of genes that play important roles in subsequent development. Whenever Dorsal's nuclear concentration exceeds a lower threshold, it *represses* the transcription of two different genes. If the concentration of Dorsal in the various cell nuclei is arranged as a gradient, each of these pairs of genes will subsequently be expressed on a different side of the embryo.

The formation of the nuclear concentration gradient of Dorsal protein is, however, entirely different from the formation of the Bicoid gradient. Overall,



GRADIENTS that have been altered misdirect development. If *bicoid* RNA is added to the posterior end of an egg (*left*), a second head and thorax start to develop there. Eggs engineered to produce a hybrid of *oskar* RNA and *bicoid* RNA (*right*) develop two abdominal ends.



CHRISTIANE NÜSSLEIN-VOLHARD

DORSAL PROTEIN gradient creates the top-bottom axis of *Drosophila* (*dark stain*). It is more concentrated in nuclei on the lower side of the embryo.

the concentration of Dorsal protein is actually level throughout the embryo. Christine W. Rushlow and Michael S. Levine of Columbia University, along with my colleague Siegfried Roth and me, have shown that what does vary along the dorsoventral axis of the embryo is the degree to which Dorsal protein is sequestered in nuclei. Close to the dorsal side of the embryo, the protein is found increasingly within the cytoplasm; on the ventral side it is found mainly within nuclei.

How does this strange gradient of Dorsal concentrated in nuclei arise? Normally, what stops Dorsal from entering nuclei is a protein called Cactus, which binds to it. On the ventral side of the embryo, however, Dorsal is released from this bound state by an activation pathway involving at least 10 proteins.

The ventral signal that starts this process originates early in egg development inside the female. Yet its effect—the importation of Dorsal to the nucleus—takes place several hours later, in embryos with rapidly dividing nuclei. Thus, the signal must be very stable. The signal's exact nature remains unclear, but it is concentrated in the specialized membrane—known as the vitelline membrane—that surrounds the egg after it is laid.

Painstaking experiments by my colleague David Stein and me and by Kathryn V. Anderson and her colleagues at the University of California at Berkeley have established that some early components of the activation pathway are produced in the mother's follicle cells, which surround the unlaid egg. Others are produced in the egg cell and then deposited either in the egg's cytoplasm or in its cell membrane or secreted into the space surrounding the egg.

Initially, the protein components of this pathway are evenly distributed, each in its proper compartment. Then the signal, which identifies the ventral side, becomes active. This signal seems to determine the Dorsal gradient by triggering a cascade of interactions among the proteins of the activation pathway; the cascade conveys into the egg the information about which side will be ventral.

This message relay system probably relies on gradients of its own. It seems likely that a true gradient first appears in the space surrounding the egg cell, because large proteins can easily diffuse through this region. The gradient signal is thought to cause graded activation of a receptor molecule in the egg's cell membrane; that is, the receptors may become either more or less active depending on how ventral their position is. The receptors could then transmit a similarly graded signal into the egg cytoplasm, and so on.

Thus, the signal that initiates the formation of the embryo's dorsoventral pattern circumvents the obstacle to diffusion. In order to do this, it relies on a message relay system that, through a variety of protein molecules, carries the gradient information from one compartment to another. (A similar mechanism for carrying a signal across the egg cell membrane operates in the terminal pathway, which is the system that controls structures at both ends of the antero-posterior axis.) In this manner, signals from outside an egg, where a gradient can easily form by diffusion, can be transmitted to the inside. The result is the graded importation into the nuclei of a protein that was initially evenly distributed.

Patterns in Common

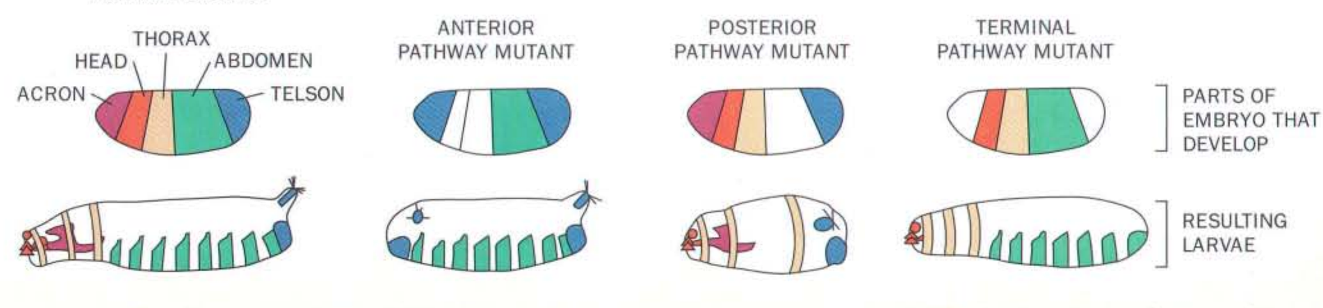
What conclusions can we draw from these investigations? Before gradients were identified, biologists believed that morphogens might constitute a special class of molecule with unique properties. This is clearly not the case. In the early *Drosophila* embryo, many "ordinary" proteins that can serve different biochemical functions can convey positional information.

In some instances, such as the process determining the dorsoventral pattern, a gradient arises first by diffusion and is then copied down a molecular chain of command by activation of successive proteins. In other cases, gradients have inhibitory effects. The Nanos gradient, for example, represses the cell's use of one type of evenly distributed mRNA, thereby creating a gradient of the opposite orientation.

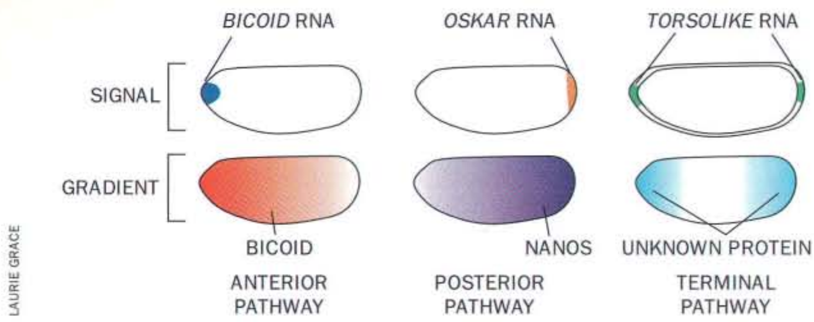
In all the pathways so far investigated, the final result is a gradient of a morphogen that functions principally as a transcription factor, initiating or suppressing the transcription of one or more target genes in a concentration-dependent manner. These gradients are sometimes quite shallow: Bicoid and Dorsal decline in concentration only slowly along the length of the embryo. Yet they somehow cause the protein products of their target genes to have sharp cutoff points. How can this happen?

One way this might occur is if several molecules—either different ones or mul-

GENES THAT DISRUPT early *Drosophila* development fall into four groups affecting different pathways. Three of the groups affect the long axis. For each, characteristic parts of the embryo fail to develop. Another group (*not shown*) affects the dorsoventral axis.



LAURIE GRACE



THREE PATHWAYS of the long axis create signals of *bicoid, oskar* and *torsolike* RNA. They give rise to four protein gradients that start subdividing the embryo.

tiple copies of the same one—cooperated to bring about transcription. The dynamics often result in a steep dependence on the concentration of one or more of the components. It is noteworthy, then, that genes activated by Bicoid or Dorsal proteins contain multiple adjacent binding sites, often for different transcription factors that may modulate the genes' activity.

Some morphogenetic gradients apparently yield but a single effect: if the concentration of the morphogen in a particular place is above a critical threshold, a target gene is activated; otherwise, it is not. In other cases, different concentrations of morphogen elicit different responses, and it is this type of gradient that is most important for providing an increase in the complexity of the developing organism.

Although each morphogenetic gradient seems to control only a few target genes directly, interactions between cofactor molecules that affect transcription can radically change responses to the gradients. These mechanisms of combinatorial regulation open the way to the formation of patterns of great complexity from an initially simple system. Proteins acting as cofactors can modify a morphogen's affinity for a gene's promoter region, thus shifting a critical threshold up or down. A cofactor might even turn an activating transcription factor into a repressor. The potential for creating complex patterns becomes apparent when one considers that the cofactors may themselves be distributed in a graded fashion.

Superposing several gradients onto an embryonic region can subdivide it even more and generate additional complexity. The three pathways that define the antero-posterior axis of the *Drosophila* embryo together give rise to four separate and independent gradients (the terminal pathway produces two gradients, of an unknown protein). Each gradient has one or two thresholds. At least seven regions are thus defined by a unique combination of target gene expression. At the anterior end, where the gradient of the as yet unidentified terminal protein and the Bicoid gradient overlap, the combination leads to the development of the foremost extreme of *Drosophila,* a part of the head. The gradient of the unknown protein acting alone, in contrast, produces the structures of the opposite end, at the tip of the abdomen.

Combinatorial regulation as a principle of pattern formation is even more apparent later in fruit-fly development. For example, the gradients of transcription factors along the long axis of the embryo affect genes that, in most cases, encode other transcription factors. Those secondary factors, in turn, diffuse out into gradients of their own. At various threshold concentrations, each factor acts on its own gene targets; sometimes these thresholds are altered by other transcription factors with overlapping spheres of influence.

Concentration dependence and combinatorial regulation together open up a versatile repertoire of pattern-forming mechanisms that can realize the designs encoded in genes. In *Drosophila,* the initial patterns generate transverse stripes of gene expression covering the part of the egg to be segmented in the larva.

This pattern in turn directs the formation of an even more finely striped pattern, which then directly determines the characteristics of each segment in the embryo. As soon as the embryo partitions itself into cells, transcription factors can no longer diffuse through the cell layers. The later steps of pattern refinement therefore rely on signaling between neighboring cells, probably with special mechanisms carrying signals across cell membranes.

Many more details remain to be discovered before we have a complete picture of how the *Drosophila* embryo develops. Yet I believe we have now uncovered some of the principal features. This accomplishment can illuminate much of zoology, because one great surprise of the past five years has been the discovery that very similar basic mechanisms, involving similar genes and transcription factors, operate in early development throughout the animal kingdom.

Basic research on a good model system has thus led to powerful insights that might one day help us understand human development. What these insights have already provided is a satisfying answer to one of the most profound questions in nature—how complexity arises from initial simplicity. SA

The Author

CHRISTIANE NÜSSLEIN-VOLHARD started her academic career studying biochemistry and gene transcription in bacteria. She turned to *Drosophila* at the University of Basel in the mid-1970s, where she initiated the research program described in this article. In 1978 she and Eric Wieschaus became group leaders in the European Molecular Biology Laboratory in Heidelberg, Germany, where the two studied genes affecting embryonic pattern formation. For the past 10 years, Nüsslein-Volhard has been director of the genetics division of the Max Planck Institute for Developmental Biology in Tübingen. She is the recipient of several scientific awards and last year shared with Wieschaus and the *Drosophila* geneticist Edward B. Lewis the Nobel Prize for Physiology or Medicine.

Further Reading

MUTATIONS AFFECTING SEGMENT NUMBER AND POLARITY IN *DROSOPHILA*. C. Nüsslein-Volhard and E. Wieschaus in *Nature,* Vol. 287, pages 795–799; October 30, 1980.

THE MAKING OF A FLY; THE GENETICS OF ANIMAL DESIGN. Peter A. Lawrence. Blackwell Science, 1992.

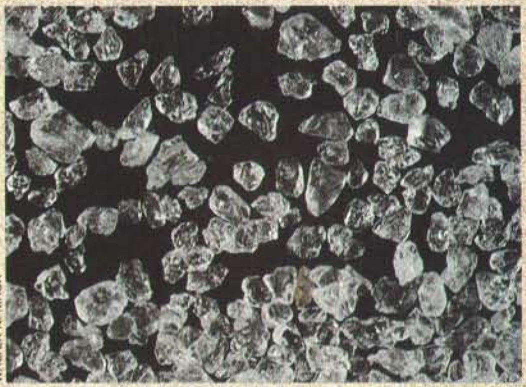
THE ORIGIN OF PATTERN AND POLARITY IN THE DROSOPHILA EMBRYO. Daniel St. Johnston and Christiane Nüsslein-Volhard in *Cell,* Vol. 68, No. 2, pages 201–209; January 24, 1992.

THE DEVELOPMENT OF DROSOPHILA MELANOGASTER. Edited by Michael Bate. Cold Spring Harbor Laboratory Press, 1993.

Sands of the World

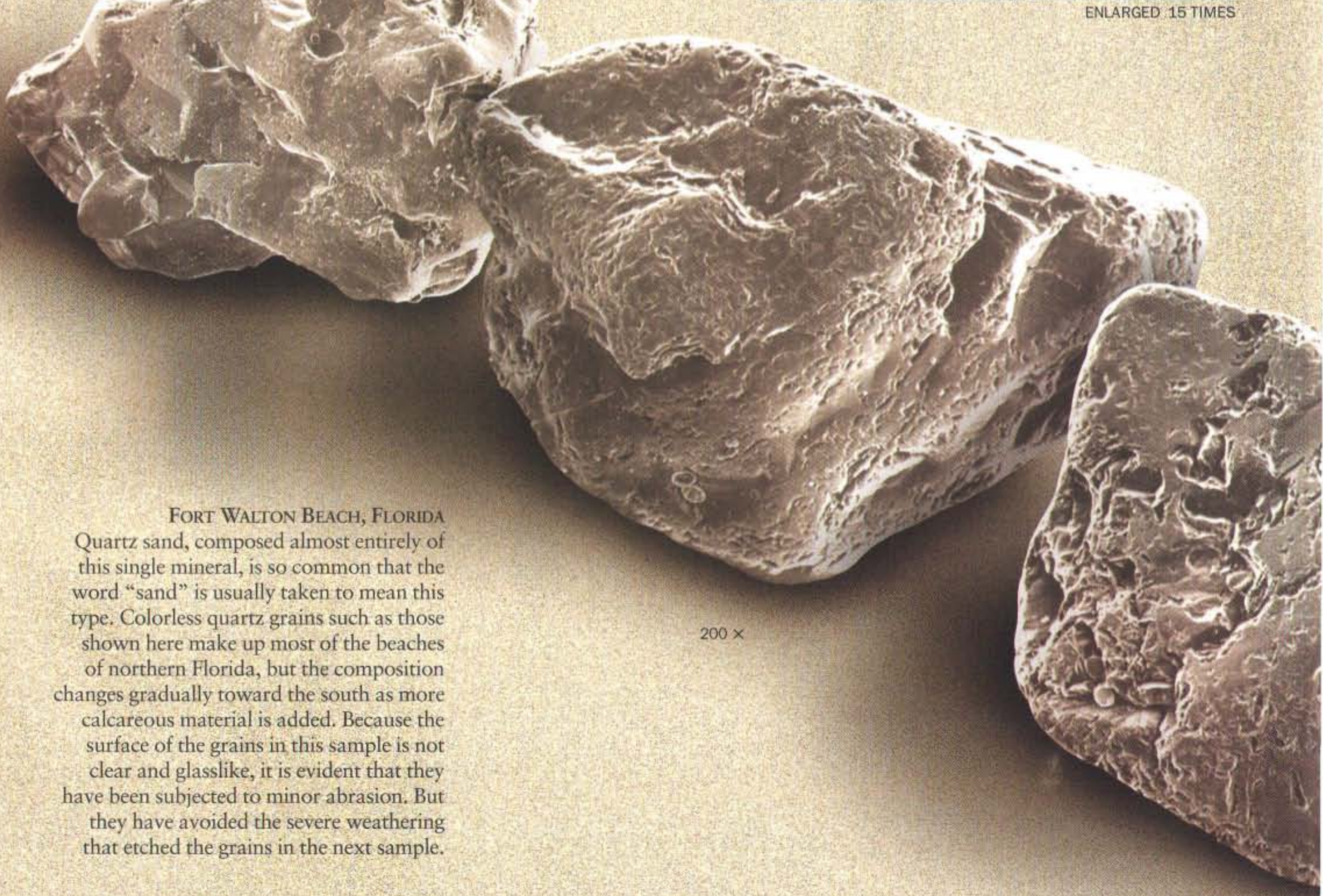
One of the most common elements on the earth's surface, sand is also one of the most various

by Walter N. Mack and Elizabeth A. Leistikow



WALTER N. MACK

ENLARGED 15 TIMES



ANDREW PAUL LEONARD *APL Microscopic*; DIGITAL MANIPULATION BY LAURIE GRACE

FORT WALTON BEACH, FLORIDA
Quartz sand, composed almost entirely of this single mineral, is so common that the word "sand" is usually taken to mean this type. Colorless quartz grains such as those shown here make up most of the beaches of northern Florida, but the composition changes gradually toward the south as more calcareous material is added. Because the surface of the grains in this sample is not clear and glasslike, it is evident that they have been subjected to minor abrasion. But they have avoided the severe weathering that etched the grains in the next sample.

TO SEE A WORLD IN A GRAIN OF SAND
AND A HEAVEN IN A WILD FLOWER,
HOLD INFINITY IN THE PALM OF YOUR HAND
AND ETERNITY IN AN HOUR.

—Auguries of Innocence, William Blake

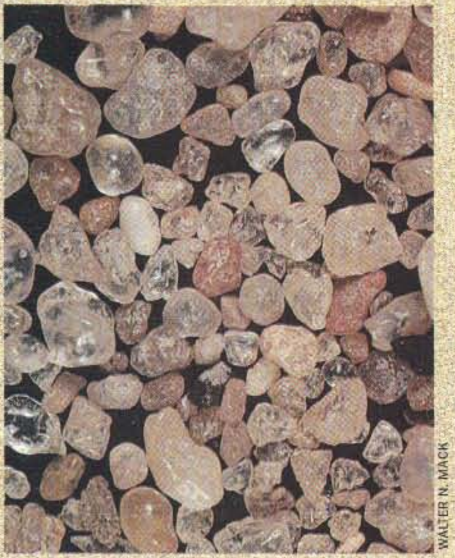
When we pick up a handful of sand from the beach and watch it sift through our fingers, we are seeing the product of millions of years of geologic history. Much of this history can be uncovered by examining the particles under magnification, where they give up the secrets of their origin and subsequent travels.

Most sand starts life in mountainous areas as continental rock—primarily as quartz and feldspar. Mechanical breakdown (by the movement of glaciers; by cycles of freezing and thawing) produces boulders and pebbles. Then chemical assault (by vegetation and rain) combines with mechanical disintegration to eat away at these boulders and pebbles, eventually giving birth to individual grains. Geologists define sand as rock fragments having a diameter between 0.05 and two millimeters; larger particles are classified as gravel, smaller as silt.

Following birth, the grains are washed downhill into a streambed. There they roll and bounce along the bottom, accumulating now in an eddying pool, now in the lee of a boulder. Years may pass before the next step of their journey, but at last they leave the mountains by way of a river. Some of the river's sand reaches the shore; some is deposited along the way. A medium-size river will take something like a million years to move its sandy deposits 100 miles downstream. In the process, chemicals in the water polish many grains to a high gloss.

Wind as well as water plays a part in distributing sand. Wherever vegetation is meager, wind sets the particles in motion. They bump and wiggle along, sometimes blown a foot or so above the ground. Grains that are transported by the wind do not become polished but take on an opaque and frosted appearance.

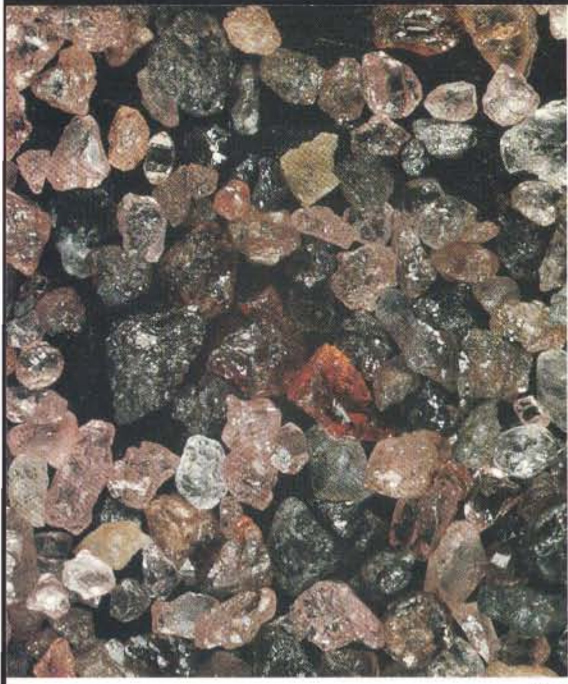
Not all sandy beaches originate as rock fragments washed and blown down from the mountains. Some beaches are composed of particles of limestone that have formed in or near the sea. And where the water is warm and the biological activity is great, beaches may consist in part or entirely of fragments of marine invertebrate animals. These are the calcareous beaches, and their "sand" grains are, by far, the most interesting to examine microscopically, because they represent some of nature's most colorful and delicate works of art.



WALTER N. MACK

7×

SAHARA DESERT, BETWEEN CAIRO AND ALEXANDRIA This sand displays telltale signs of wind erosion. The dull, opaque surfaces reflect the buffeting grains receive when they are transported by wind rather than by water. The wind-borne particles are roughed up more because they lack the buoyancy and buffering provided by a watery medium; their contact with other particles thus subjects them to more abrasion. The speed of the wind also exposes the grains to more punishment. Another difference is evident as well: desert sands tend to have a wider assortment of grain sizes. Water sifts its sediments more selectively than air does, depositing particles of similar size close together.



WALTER N. MACK

23×

NORTH BEACH, HAMPTON, NEW HAMPSHIRE
As the North American ice sheet receded, it dumped vast amounts of debris along the rugged shoreline of the northeastern U.S. This specimen provides a sampling of those deposits—a mixture of quartz (*colorless grains*), feldspar (*pink and amber*), and opaque igneous minerals (*black*).

WALTER N. MACK

6.5×

PUNALUU, HAWAII
The sand of Hawaii's famous black beaches is obsidian—volcanic glass created by magma that flowed into the sea, where it cooled so rapidly that it vitrified. Water and waves worked on broken bits of the glass, eventually reducing it to fine black sand.

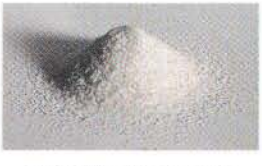
WALTER N. MACK

30×

SOUTH SHORE OF LAKE SUPERIOR, MICHIGAN
Many beaches show intriguing black streaks at the waterline. The streaks appear to be composed of organic debris or oil-soaked sand, but they are actually made of particles of magnetite. Heavier than the surrounding grains, these hard, magnetic particles are left at the water's edge as the waves toss the lighter quartz fragments higher up on the beach. (In this sample, the quartz grains are pink; the deep red may be garnet.) Twelfth-century navigators placed magnetite, or lodestone, as they called it, in a hollow reed; by carefully floating the reed in a bowl of water, they obtained a north-south bearing with this crude form of compass.

NORTHERN LIGHT LAKE, ONTARIO, CANADA
Not all black beaches are obsidian or magnetic sand. The beach on Northern Light Lake, for example, is a deposit of fine crystals of hornblende (a complex silicate mineral). Both the lake and its hornblende beach were left behind by the North American ice sheet.

Sand Samples without Magnification



FORT WALTON BEACH
Florida

SAHARA DESERT
between Cairo and Alexandria

NORTH BEACH
Hampton, New Hampshire

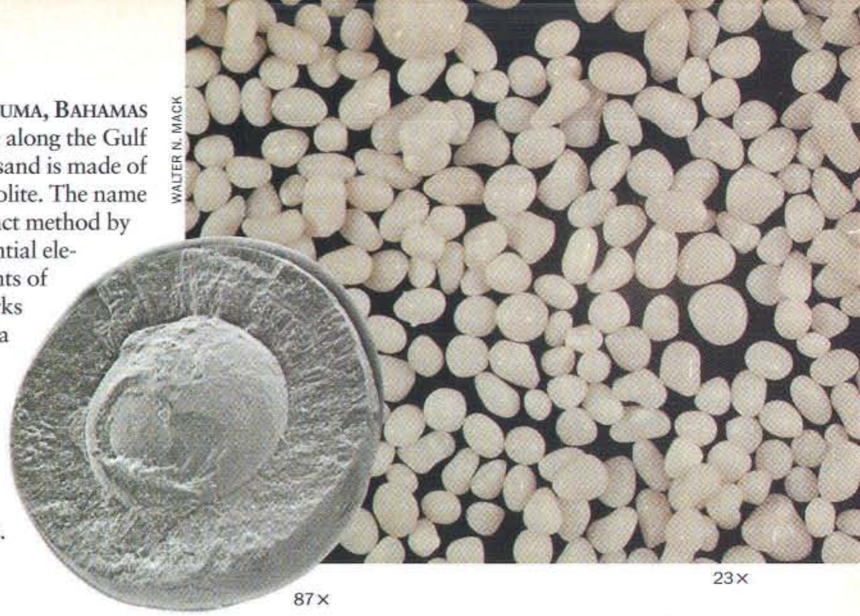
PUNALUU
Hawaii

Hawksbill Cay, Exuma, Bahamas

Some of the most photogenic, glistening white beaches lie along the Gulf of Mexico and on the islands of the Bahamas. Their sand is made of smooth, porcelain-hard particles of limestone called oolite. The name comes from the Greek, meaning "egg stone." The exact method by which these tiny "eggs" are formed is unknown. One essential element is shallow water that contains precipitating amounts of calcium and magnesium carbonate. Another is ripple marks on the sea bottom, which enable wave action to rotate a particle of clay or fine sand, on which concentric layers of carbonates then form.

Inset: Seen under a scanning electron microscope, a fractured grain of oolite discloses the lamination around the core of the particle (which is itself not visible)—layer after layer of calcium and magnesium carbonate.

WALTER N. MACK



87×

23×

WALTER N. MACK

18×

Silver Sands Beach, Grand Bahama

The "sand" shown here is almost exclusively coral rubble from the nearby reefs. There are, in addition, two cigar-shaped spicules, a cross section of a marine worm tube, a gastropod shell and two large, round red foraminifers.

WALTER N. MACK

23×

Indian Key, Florida

All the coral and shells in this sample have lost their gloss, leaving the exteriors dull, chalky and pitted. This condition is sometimes seen on the white beaches of the tropical Florida Keys, attesting to the decay of calcareous beach material. Warm seawater, direct sunlight and abundant freshwater from rain can conspire to take back into the sea the carbonates of the dead plant and animal skeletons. We see here four fusiform and one globular gastropod shell and the remains of at least two bivalve shells, all in the process of being reclaimed by the ocean.

WALTER N. MACK

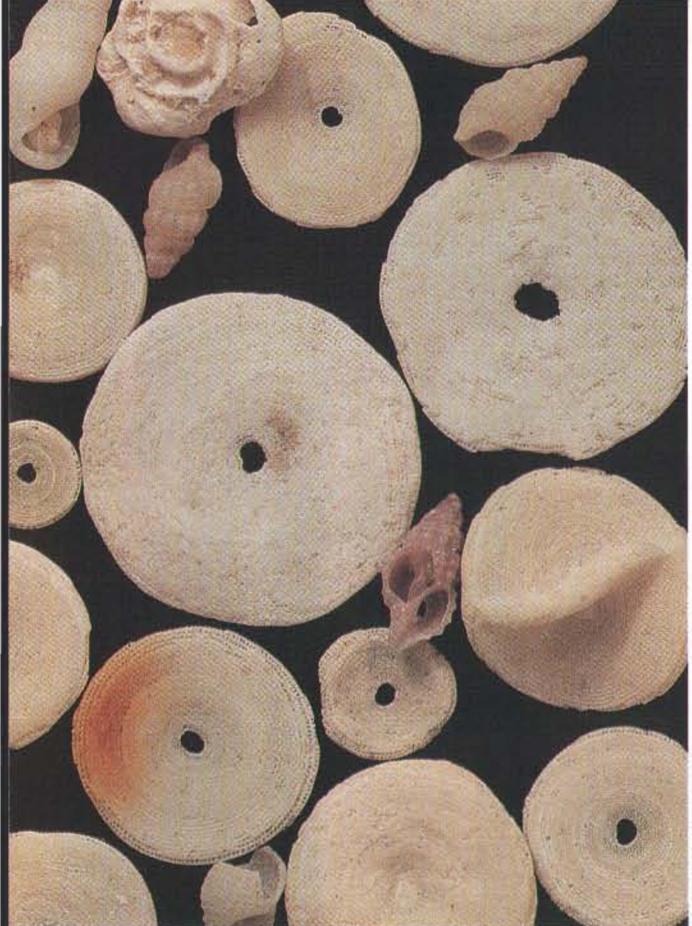
5×

South Shore of Lake Superior
Michigan

Northern Light Lake
Ontario, Canada

Hawksbill Cay
Exuma, Bahamas

Silver Sands Beach
Grand Bahama



WALTER N. MACK

6.5×

WALTER N. MACK

21.5×

LIFUKA ISLAND, HAAPAI GROUP, TONGA, SOUTHWEST PACIFIC
The remains of crinoids make up part of the sand on some islands of the South Pacific. Originally thought to be plants (their common name is sea lily), these animals have a long stem that consists of a series of calcified, wheellike plates. After the animal dies and the soft tissue decays, the stem becomes separated, and the stony disks fall in large numbers to the ocean bottom. Some find their way into the calcareous deposits of a beach. The disks vary in shape, depending on the species from which they come. The periphery of several pictured here has been eroded, disclosing the complex compartmentation of the interior structure.

WALTER N. MACK

INDIAN KEY
Florida

LIFUKA ISLAND, HAAPAI GROUP
Tonga, Southwest Pacific

SEVEN MILE BEACH
Dongara, Australia

TAKETOMI SHIMA
Ryukyu Islands, Japan

DAN WAGNER

WALTER N. MACK

17×

SAINT-TROPEZ, FRENCH RIVIERA

Just off Saint-Tropez the reefs support many interesting animals whose shells are tossed onto the beach by the waves. In this sample, the conical gastropod shells display their complex beauty. One has had holes drilled through it by a hungry predator; others reveal debris firmly wedged into the aperture. The long, curved tubular shell belonged to a mollusk in the genus *Caecum*. This creature begins life as a miniature, normal coiled snail but then grows in a single direction only. Below it lies the white, slightly abraded horn of a marine ram (*Skeneopsis planorbis*). Near the center is a large black and gold mica crystal; the reddish brown rod above it is a sponge or sea-urchin spine.

SEVEN MILE BEACH, DONGARA, AUSTRALIA

Just off Seven Mile Beach, in the Geelvink Channel, lies a shallow continental shelf teeming with life from the Indian Ocean. Many small corals and shells are evident in this photograph; however, the most prominent objects are the three-axial, iciclelike sponge spicules and the very immature globular and discoid gastropod and bivalve shells.

15×

TAKETOMI SHIMA, RYUKYU ISLANDS, JAPAN

Some of the southern Japanese islands are known for their beautiful star sand. Star sand grains are the shells, called tests, of foraminifers, microscopic, single-celled animals that abound in the world's oceans. The shells are the only means of classifying the animals. This sample contains primarily *Baculogypsina sphaerulata*. There is also a single round foraminifer test, from *Amphistegina madagascariensis* (*upper right*), and a single glass spiral gastropod shell (*left of center*).

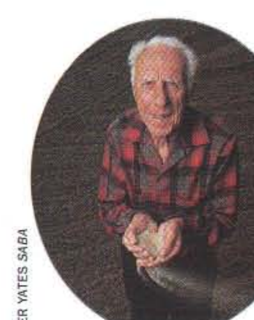
6×

WALTER N. MACK

SEAFORD, ENGLAND

The city of Seaford is in southern England, on the English Channel, where tidal currents are strong and the water quite cold. Nevertheless, a sample of the beach sand discloses a surprising amount of animal life and several other noteworthy features. The flat blue and brown objects are bivalve fragments; a single white gastropod shell (*bottom left*) exhibits the growth of two body whorls. The three aggregations of sand grains are all neatly cemented together (*far left, just above center* and *far right*). The usual quartz grains are frosted, yet one grain (*far right*) is angular and the surface unscratched, as if it were recently added to the more weathered bits.

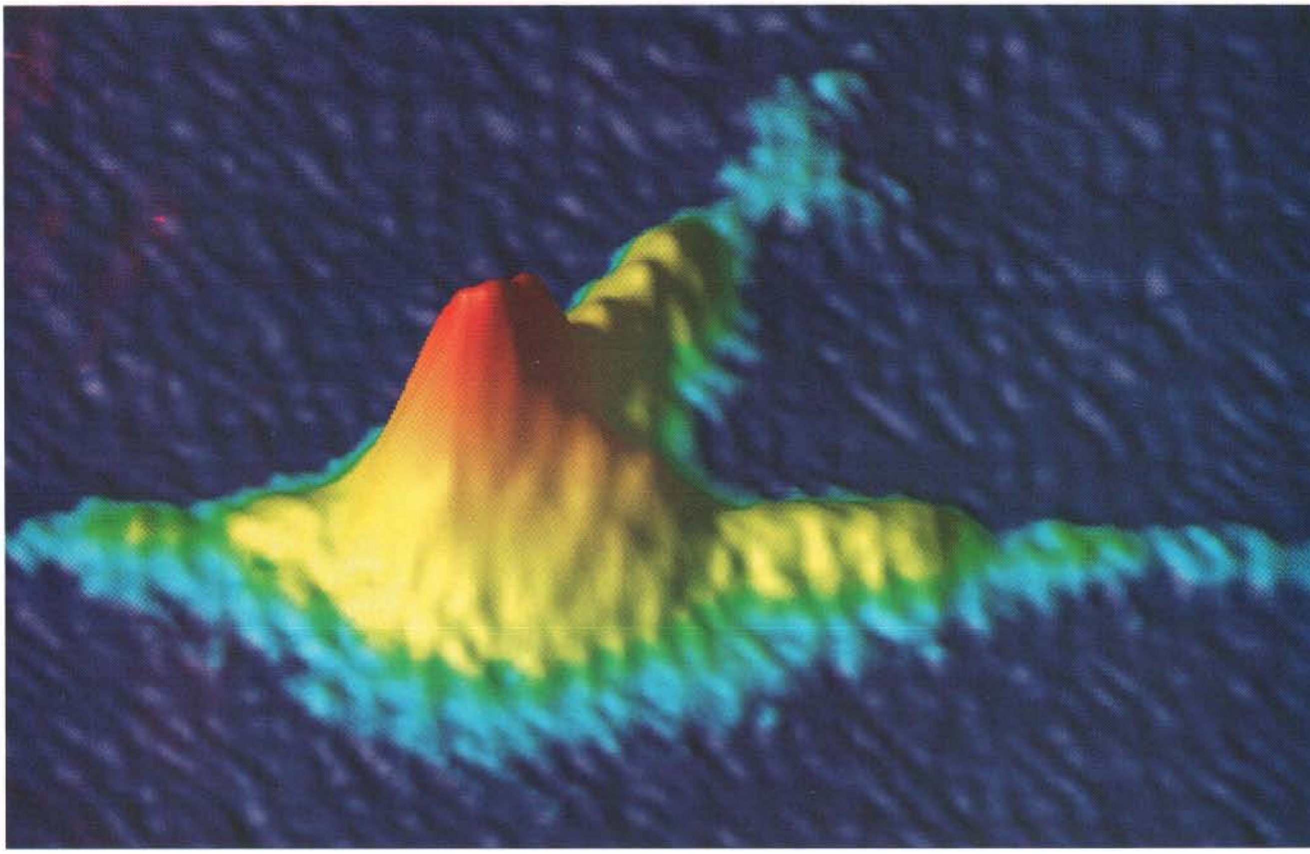
The Authors



PETER YATES SABA

REDUCED 32 TIMES

WALTER N. MACK and ELIZABETH A. LEISTIKOW have long shared an interest in sand. Trained as a microbiologist, Mack (*left*) received a Ph.D. from the University of California, San Francisco, in 1947. He taught at Michigan State University for more than 30 years before retiring in 1977. His avocational fascination with sand began early in his career, while he was on vacation at a Lake Michigan beach. Since then, he has collected—with the aid of friends and acquaintances—sand from all corners of the earth, put it under his microscope, photographed it and studied it. Leistikow, a former student of Mack's, holds an M.D. and a Ph.D. from Wayne State University. She is with FSH/Mayo Health Systems in La Crosse, Wis.



CLIFFORD A. PICKOVER *IBM Research*

Probing High-Temperature Superconductivity

Recent experiments exploiting subtle quantum effects yield important clues about why some ceramics conduct electricity without resistance

by John R. Kirtley and Chang C. Tsuei

One of the most memorable sessions in physics occurred on March 18, 1987, during a meeting of the American Physical Society. Hastily arranged to accommodate a huge number of postdeadline talks, the gathering in the Hilton Hotel in New York City drew 2,000 physicists. Crammed into a ballroom, with many others spilling into the hallway, they struggled to give and to hear five-minute briefings on the latest ideas and investigations. Dubbed the Woodstock of physics, the session started at 7:30 P.M. and continued until 3 A.M.—even at that hour the excitement was still evident.

The cause of all this academic commotion was the announcement of high-temperature superconductivity. Late in 1986 J. Georg Bednorz and K. Alexander Müller of the IBM Research Laboratory in Zurich had reported that a ceramic called lanthanum barium copper oxide lost all electrical resistance when it was cooled to only –238 degrees Celsius, or 35 kelvins (degrees above absolute zero). Although that temperature is still quite frigid, it was nonetheless more than 10 degrees better than the best conventional superconductors, which are made from metals or alloys. Soon thereafter critical temperatures above 90 kelvins were reported and confirmed, and rumors of superconductivity at 130 kelvins and 240 kelvins abounded. If a material could be found that superconducted at room temperature (300 kelvins or so), it would very likely initiate a revolution in modern society.

At the March meeting, physicists had submitted papers on theory and measurements of the new superconductors. The frantic nature of the session was caused not only by the dream of room-temperature superconductivity but also, in part, by fear. Some worried that these copper oxide materials, often referred to as cuprates, would be understood before each researcher could make a meaningful contribution. After all, superconductivity is the single topic for which the most Nobel Prizes have been awarded (five so far). A correct theory would surely garner another for the field.

In hindsight, the participants had little to worry about. In the nine years since, thousands of scientists around the world have devoted millions of hours to the question of why and how the cuprates superconduct at such high temperatures. Although neither question has been satisfactorily answered, physicists have made much progress. Recent experiments have shown that the cuprates are fundamentally different from conventional superconductors and have helped define the field of competing theories. They indicate that a radical mechanism is still in the running: magnetic fluctuations in the atoms that make up the conducting medium.

Conduction by Cooper Pairs

The idea that magnetism could lie at the heart of the matter contrasts sharply with our well-established understanding of the mechanism of conventional, low-temperature superconductivity. Conventional superconductivity happens when the electrons combine to form so-called Cooper pairs (after Leon N. Cooper, then at the University of Illinois, the physicist who first introduced the concept). Unlike single electrons, Cooper pairs do not bump into one another or scatter off the imperfections in the conducting medium. Hence, they face no resistance to their forward motion. An electric current in a superconductor will flow without voltage and, once set in motion in a loop, will persist forever—so long as the substance is kept cooled to below the critical temperature.

It is remarkable that electrons in metals can pair. After all, they are all negatively charged and therefore normally repel one another. During the 1950s, Cooper, along with his colleagues John Bardeen and J. Robert Schrieffer, came up with an explanation. The BCS theory, as it is called in honor of its originators, states that electrons in conventional superconductors overcome their mutual repulsion in two ways. First, some of the negative charge is blocked off. Such a "screening" effect, which stems from the motions of other electrons, reduces the repelling force between the electrons in a Cooper pair.

Second, and more important, an intermediary can act to bring electrons together despite their repulsion. This matchmaker consists of the positive ions that make up the metal. (The neutral atoms become positive ions after they donate electrons for conduction.) A moving electron can slightly shift the positions of the ions as it passes by. These distortions, known as phonons, create small positive patches that attract other electrons. A typical analogy is that of bowling balls on a bed. One bowling ball will distort the springs in the mattress. This distortion will tend to draw in a second bowling ball.

This analogy goes only so far, because the electron bowling balls strongly repel one another. A better one came from Bardeen. He referred to a closely packed crowd that has stormed a football field. The Cooper pairs can then be thought of as couples in the melee, desperately trying to stay together. Once set in motion, such a crowd is hard to stop, because to stop one person in the group, you must stop many others. The crowd members will flow around obstacles, such as goalposts, with little disruption.

Physicists like to say that electrons remain paired by exchanging phonons, just as rugby players can pair by passing the ball between them to avoid being tackled as they move downfield. The phonon-pairing mechanism, when embodied in the BCS theory, works extremely well for explaining superconductivity in conventional materials.

But researchers generally believe that the conventional BCS picture by itself cannot explain superconductivity in the copper oxides. Electrons and phonons in a BCS superconductor with a high critical temperature would interact very strongly with one another. In that case, the structure of the material would end up being distorted in such a way that the material would no longer be superconducting and usually not even conducting.

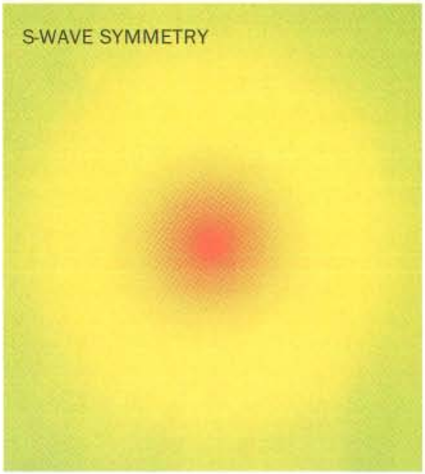
Furthermore, the BCS model relies on the electrons being much more energetic than the phonons. The electrons move much faster than the phonons do, so that the first electron has moved far away from the displaced ion when the second electron arrives. That distance between electrons reduces the effects of the negative repulsion between them. But in the cuprates the electrons and phonons would move at similar rates, so that there would not be much distance between electrons making up the pairs.

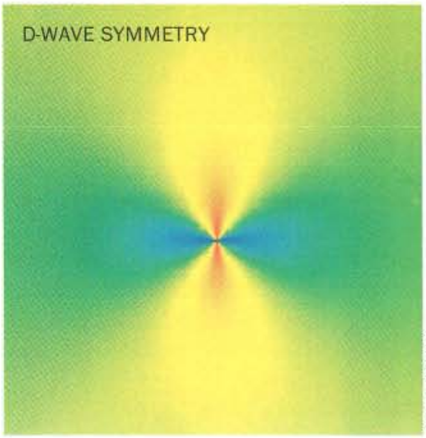
(We mention electrons as the carriers of electricity, but the truth of the matter is that, in most cuprates, the carriers of electric current are the "holes," or the positively charged spaces left by electrons. Such holes arise when extra atoms called dopants are added to the substance to soak up some electrons. For the rest of this article, we will use the word "carriers" rather than "electrons" when describing the components making up the Cooper pairs.)

Because of the difficulty in explaining high critical temperatures using phonons,

IBM RESEARCH

MAGNETIC FLUX threads through the intersection of three boundaries in an unpatterned superconducting film (*opposite page*). This so-called half-flux quantum indicates that radical physics may underlie high-temperature superconductivity, which someday may lead to such applications as the levitation of trains. Here a magnet rides above an yttrium-based superconductor (*right*).





JOHN R. KIRTLEY

TWO TYPES OF SYMMETRY of the superconducting wave function are s-wave and d-wave. In the s-wave condition, one member of a Cooper pair is located in the spherical area around its partner. For d-wave symmetry, the partner lies somewhere in one of four lobes, which are negative (*blue*) or positive (*red yellow*).

many other pairing intermediates—the types of "balls" passed between "rugby players"—have been proposed. They include excitons, in which the charge carriers cause localized motions of the electrical charge cloud around them; plasmons, where the carriers excite collective motions of the surrounding charge cloud; and polarons, in which the carriers produce large displacements of the ions and other charge carriers through which they move. Other ideas treat each charge as two distinct particles that can hop between layers in the material.

Unfortunately, the cuprates are complicated materials. They consist of several layers that insulate the copper oxide layer, which is the main conducting zone. No tests that unambiguously distinguish between pairing mechanisms have been formulated or, at least, agreed on. There has, however, been progress on one potentially powerful feature of the superconducting state: its symmetry. This feature may serve as a litmus test of proposed pairing mechanisms.

Wave Function Symmetry

Symmetry refers to the form taken by the superconducting state's mathematical description, or wave function. Among the other features that it characterizes, the wave function reveals how the two carriers making up the Cooper pair move with respect to each other. It indicates the probability of finding one partner as a function of its position relative to its mate.

The Cooper pairs of conventional superconductors take on the most symmetric wave function possible: spherical, or s-wave, symmetry. That is, the chance of finding one carrier in a Cooper pair given the position of the other falls off at the same exponential rate in all directions in space. If we plotted the wave function, keeping one member of the Cooper pair at the center, the probability of finding its partner would appear as a sphere around the center.

The next most highly symmetric state for the cuprates is the d state. Plotted, it would appear as four lobes lying in a plane, like a four-leaf clover. Each lobe represents a likely position of one member of the Cooper pair with respect to its partner. D symmetry also means that the Cooper pair members are not so close to each other that their mutual repulsion interferes with their coupling.

How would identifying the symmetry of the superconducting state help to define the mechanisms that create Cooper pairs? It turns out that some of the proposed mechanisms produce a distinct symmetry. Until a few years ago, most theorists fell into one of two camps, roughly speaking. One group favored mechanisms that yielded s-wave symmetry states; most, but not all, of these theories tended to be modifications of the BCS phonon-mediated theory. Other theorists believed in mechanisms that generated d-wave symmetry states. Such proposals tended to be a rather radical reworking of the underlying physics.

Perhaps the most dominant theory that contains d-wave symmetry has been the spin-wave model. Douglas J. Scalapino of the University of California at Santa Barbara and David Pines of the University of Illinois are the main champions of this theory. In the scenario, a moving charge can disrupt the orientation of the spins of the atoms that make up the superconducting medium. In effect, the carrier leaves a magnetic disturbance (a spin wave) in its wake. This wake pulls in a second carrier, so that the two form a Cooper pair. The spin waves are short-lived, so they are often called spin fluctuations.

Many physicists had thought that the symmetry of the superconducting state might indicate the correct theory. But matters proved more complex, and theorists discovered in the past couple of years that different mechanisms could produce the same symmetry. Hence, identifying the symmetry will not in itself nail down the mechanism. More accurately, successful identification would enable theorists to refine their models. For instance, an unambiguous demonstration that the superconducting state is not d-wave symmetric would rule out spin waves as the pairing mechanism.

Hints of D Waves

One testable property of d-wave symmetry is that the Cooper pairs are more weakly bound in some directions than in others, relative to the underlying atomic lattice. As such, it might leave unpaired carriers traveling along certain directions. Investigators have conducted a number of tests for these unpaired carriers. Such tests involved how well magnetic fields penetrated the superconductor or how much heat was needed to warm the material. The results of these experiments generally favor the presence of unpaired carriers at low temperatures, but they did not convince most physicists. The tests were somewhat indirect, and the findings were often consistent with other symmetry states (including modified s-wave states).

So instead of looking for free carriers, other physicists sought to measure how the strength of the Cooper pairing varies with angle. They tried to measure this angular dependence by seeing how charge carriers become excited by high-frequency light to escape from the sample, how light sent through the material shifts in frequency and how electrons tunnel into the cuprates from other materials through thin insulating barriers.

These studies found that an angular variation to the pairing strength exists. But they failed to be the smoking-gun proof for d waves for another reason. In d-symmetry states the wave function changes from plus to minus and back again; in other words, the four lobes al-

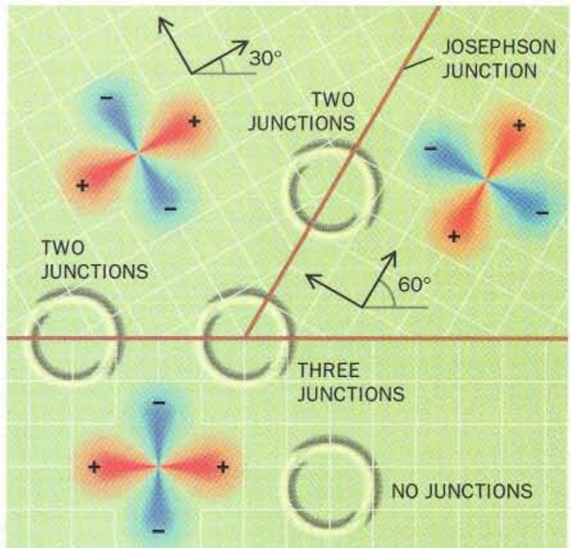
ternate in sign: two lobes are positive, and the other two negative. The tests were insensitive to the sign differences, and so the data did not constitute a convincing victory for d-wave symmetry.

There is, however, a way to see the positive and negative signs in the lobes. The method takes advantage of a well-known property: superconductors shaped into rings can trap magnetic fields in the space enclosed by the rings. The fields are trapped in discrete bundles, known as flux quanta. A single flux quantum can be viewed as having a tube-like shape. It has a total magnetic flux (magnetic field times the area enclosed by the ring) that equals a fundamental constant. (Specifically, it equals $h/2e$, where h is Planck's constant and e is the charge of the electron.) Rings of conventional superconductors, which display s-symmetry wave functions, enclose integer multiples of these flux quanta.

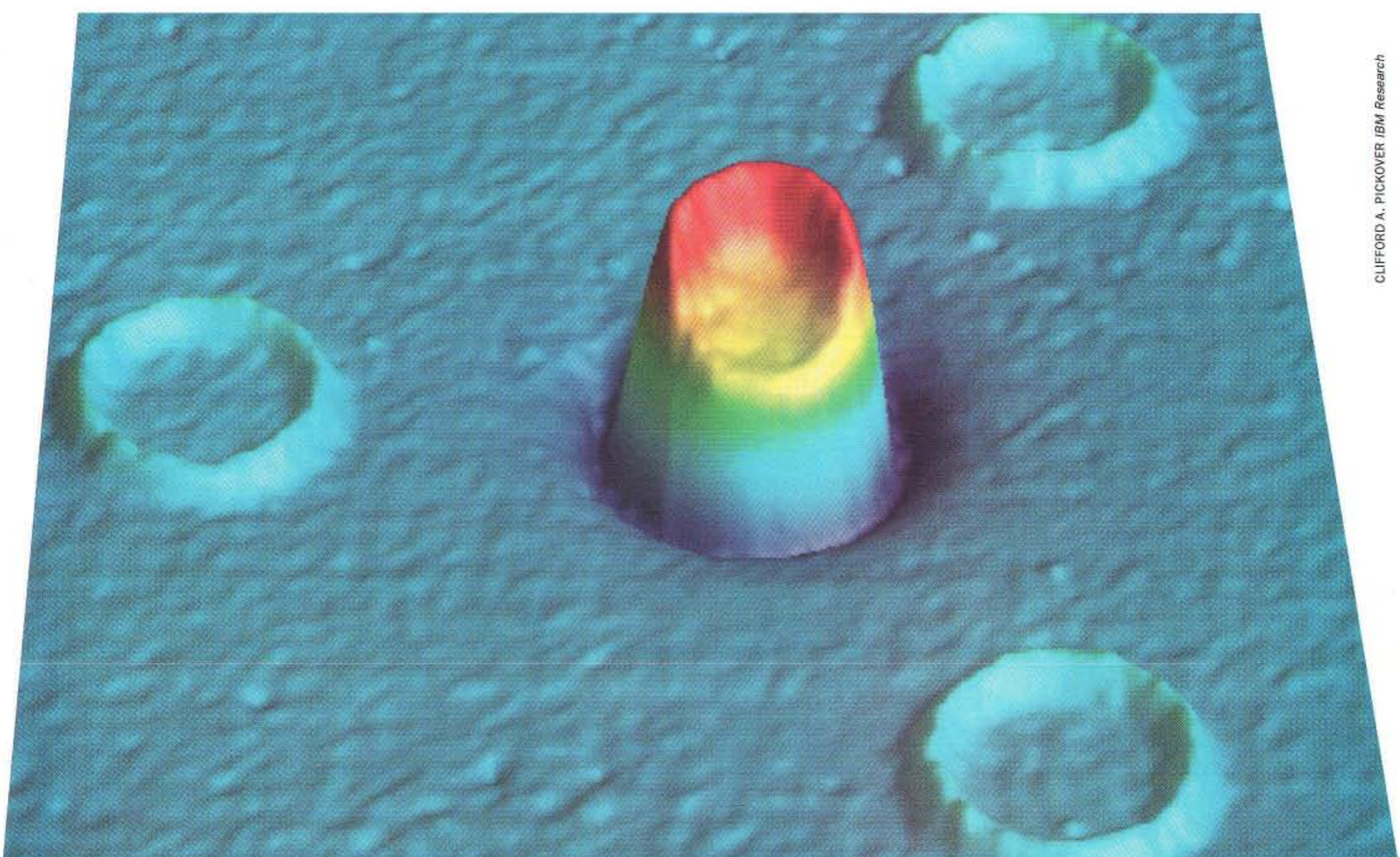
Rings made from superconductors that have a d-symmetry state, however, can quantize flux in a different manner. Based on calculations of the system's energy and magnetic flux, such rings can trap half-integer multiples of the flux quantum. It turns out that the presence or absence of half-integer flux quantization can determine whether, and in what way, the lobes change sign.

Researchers have been able to detect half-integer flux quanta only recently, even though Lev N. Bulaevskii, now at Los Alamos National Laboratory, and others predicted their existence in the late 1970s. The first experimental hints came in 1993 from magnetization studies done by Dieter Wohlleben of the University of Cologne. Later, Dave Wollman and Dale Van Harlingen of the University of Illinois reported more concrete evidence by measuring current and voltage characteristics of superconducting rings made of a single crystal of yttrium barium copper oxide (YBCO) and a thin film of lead.

Recently our group at IBM made the first direct observations and images of these half-flux quanta. We relied on spe-



"TRICRYSTAL" RING EXPERIMENT tested d-wave models (*right*). Three sections of an yttrium-based superconductor were oriented to create boundaries called Josephson junctions. Rings were etched so that one had three junctions; the others had none or two. If the superconducting state were d-wave, then a magnetic half-flux quantum should form inside the three-junction ring; the others, with even numbers of junctions, should produce no flux. The image of the half-flux quantum (*below*) proves that the d-wave theories are viable explanations.



CLIFFORD A. PICKOVER *IBM Research*

cially designed cuprate rings that are interrupted by thin layers of insulating materials. Such barriers, called Josephson junctions, are slim enough so that Cooper pairs can quantum-mechanically tunnel through them. The tunneling is referred to as the Josephson effect, after Nobel laureate Brian Josephson, who first predicted the phenomenon in 1962.

Josephson tunneling takes place only if Cooper pairs on one side of the junction are "out of phase" with those on the other. "Phase" here refers to one of the main features of a wave function (loosely speaking, it describes which part of a cycle the wave function is in). With d-wave superconductors, one can design rings with Josephson junctions that automatically change the phase of the Cooper pair circulating around the ring. This phase change is equivalent to a sign change in the wave function.

Therefore, on cooling, this built-in sign change will spontaneously generate just enough current to enclose exactly half of a magnetic flux quantum. When cooled in magnetic fields, flux values of 3/2, 5/2, 7/2 and so on, multiplied by the flux quantum, will be threaded through such a ring.

Ring around the Flux

We grew thin-film rings of the yttrium-based superconductor on a specially designed substrate in such a way that one of the rings contained three sections. Each section was misaligned by 30 degrees with its neighboring sections, so that each boundary formed a Josephson junction. If the Cooper pairs exist in a d-wave symmetry state, then the pairs completing a circuit around the ring would end up changing sign. (We do not actually know how many sign changes there are, only that there have to be an odd number for this geometry.)

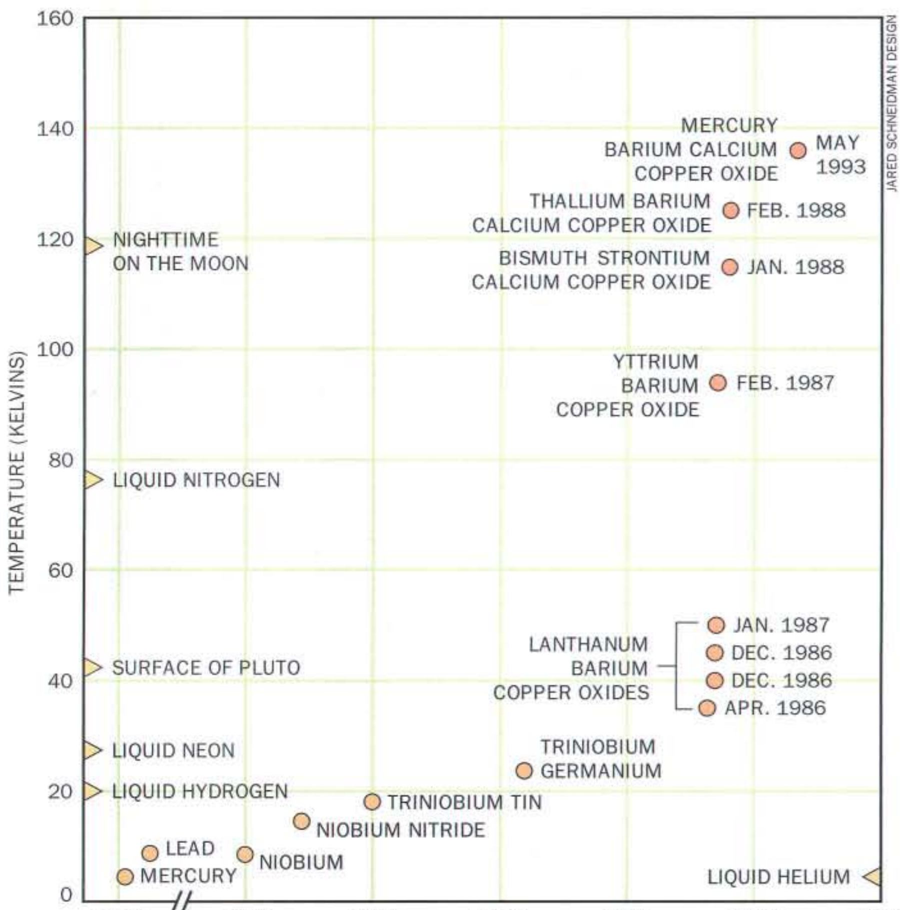
In contrast, if the material were s-wave, which does not have any sign changes, the boundaries would not matter. No sign changes would occur on completion of a circuit.

After fabricating these "tricrystal" rings, which were about 50 microns across, we cooled them to below their critical temperature. Because of their geometry, the rings were naturally unstable in terms of conduction, so a small supercurrent spontaneously developed. In a sense, the ring acts as if it has a single kink, as in a Möbius strip. The tendency of the kink to unwind sets the charge carriers in motion. We found only half-integer flux quanta in our ring: the sure sign of d-wave symmetry (s-wave would have produced none).

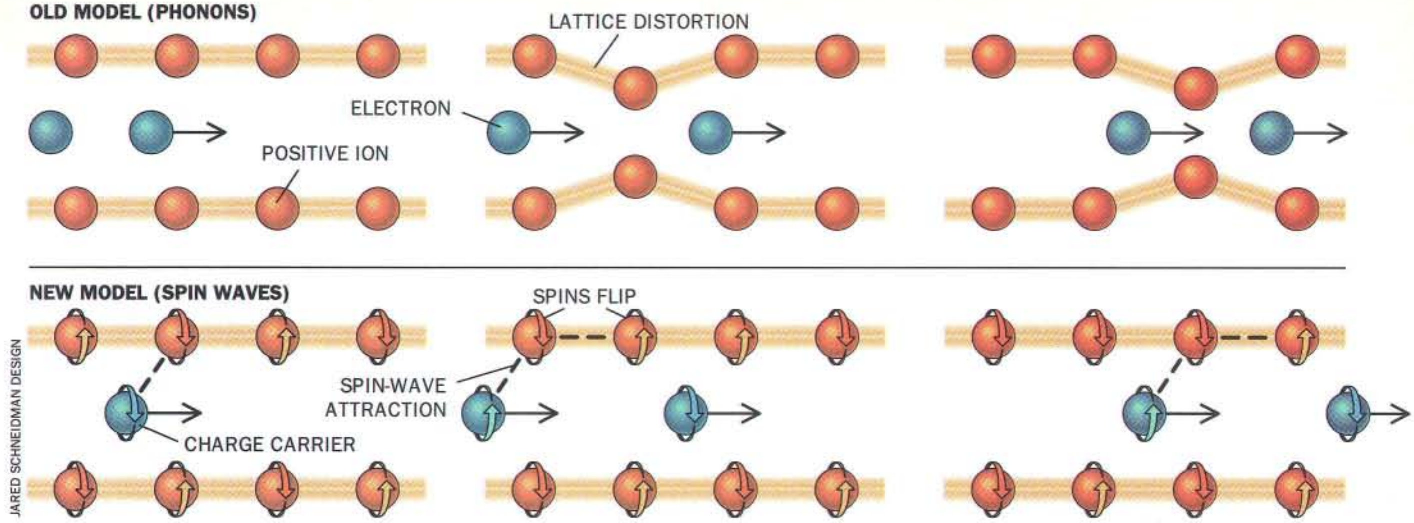
We imaged the magnetic fields trapped in the rings using a scanning superconducting quantum interference device (SQUID) microscope. The SQUID is the world's most sensitive magnetic-field sensor [see "SQUIDs," by John Clarke; SCIENTIFIC AMERICAN, August 1994]. Careful calibration of the SQUID output signal in several different ways ascertained that there was exactly a half-flux quantum in the tricrystal ring. Rings made with an even number of junctions served as controls and showed no flux quanta (because the sign flips an even number of times and thus returns to its original condition).

Our group also manipulated the conditions slightly to show that the results were in fact indicative of the underlying symmetry of the Cooper-pair wave function and not, say, caused by some other physical effects. We showed that slight changes in the geometry of the rings turned the spontaneous half-flux quantum on and off. By applying a weak magnetic field, we were also able to get the other rings to trap integer multiples of flux quanta, thus demonstrating that all the rings were indeed functioning. Experiments with unpatterned films and disks, as well as rings, also revealed the half-flux quantum effect, proving that the result derives from the intrinsic symmetry of the superconductor and not from the details of the sample geometry.

During the past year, we repeated these experiments with three other cuprates: bismuth strontium calcium copper oxide (which is in some sense more complicated than YBCO), gadolinium barium copper oxide (about as complex as YBCO), and a superconductor of thallium barium copper oxide (simpler than YBCO). We obtained the same results. Our experiments, as well as those done at the University of Illinois, the Swiss Federal Institute of Technology in Zurich and the University of Maryland, are all consistent with d-wave symmetry.



SHARPLY RISING CRITICAL TEMPERATURES in superconductors stem from the cuprate materials. None of the proposed theoretical mechanisms rule out room-temperature superconductivity, although physicists have yet to see any irrefutable signs.



OLD AND NEW MODELS of superconductivity, shown here highly simplified, rely on the pairing of charge carriers. In the old model, which explains low-temperature superconductivity, an electron produces phonons: distortions of the positive lattice ions. The phonons attract a second electron. In the spin-wave model, one of several new theories, a charge carrier disrupts the magnetic spin of one ion, which flips the spin of a neighbor and thereby attracts a second charge carrier of opposite spin.

Today there can be little doubt that the superconducting wave functions in several cuprate superconductors demonstrate d-wave symmetry. (Some experiments reported s-wave symmetry, but that might be explainable because, under certain circumstances, a cuprate may combine both types of symmetries.) This result means that, not surprisingly, the conventional BCS model used to explain superconductivity cannot be correct for the high-temperature superconductors. It also means that spin waves remain a viable pairing intermediate.

What Lies Ahead?

Unfortunately, however, nearly every other pairing mechanism that has been proposed (excitons, polarons and others) can be made to agree with these results. One just has to assume that the carriers in a Cooper pair strongly repel each other. Such an interaction favors a pairing symmetry with sign changes.

Nevertheless, symmetry tests can still do much to narrow down the field of possible pairing mechanisms. It is important to repeat these experiments with other cuprate superconductors. For example, neodymium cerium copper oxide, when doped in a way that donates electrons, seems to have s-wave symmetry. This possibility, if true, would be damaging to those who favor the spin-fluctuation mechanism, because many researchers like to think that the same mechanism applies to all the cuprate superconductors. It would also mean that the substances are even more intricate than imagined. A systematic study of the symmetry as a function of composition may help weed out invalid theories.

On a positive note for those interested in commercial prospects: most of the proposed mechanisms do not rule out a superconductor at room temperature. For instance, back-of-the-envelope calculations with the spin-wave model hint that superconductivity might be sustainable well above 20 degrees C, perhaps to several hundred degrees. Although the upper range is undoubtedly unrealistic, the possibility of a room-temperature superconductor is nonetheless a big improvement over predictions by the conventional BCS theory, which puts the limit at about –233 degrees C (40 kelvins). Nailing down the mechanism may help those trying to fabricate new superconductors and to devise practical uses for them [see "High-Temperature Superconductors," by Paul C. W. Chu; SCIENTIFIC AMERICAN, September 1995].

Certainly much more work remains before physicists can conclusively determine the pairing mechanism. But by looking for magnetic flux quanta threading through rings, researchers have a powerful new tool that can help determine why these complex materials resist explanation—but not electricity.

The Authors

JOHN R. KIRTLEY and CHANG C. TSUEI look for half-flux quanta in superconducting rings at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y., their employer since the 1970s. Kirtley, a research staff member, received his B.S. and Ph.D. from the University of California, Santa Barbara. Tsuei earned his B.S. from the National Taiwan University and his Ph.D. from the California Institute of Technology. Fellows of the American Physical Society, both have explored the fundamental electrical and magnetic properties of solids in addition to their studies of superconductivity.

Further Reading

THE CASE FOR DX2Y2 PAIRING IN THE CUPRATE SUPERCONDUCTORS. D. J. Scalapino in *Physics Reports*, Vol. 250, No. 6, pages 329–365; January 1995.

CALL THEM IRRESISTIBLE. Tim Folger in *Discover*, Vol. 16, No. 9, pages 83–91; September 1995.

DESIGN AND APPLICATIONS OF A SCANNING SQUID MICROSCOPE. J. R. Kirtley, M. B. Ketchen, C. C. Tsuei, J. Z. Sun, W. J. Gallagher, Lock-See Yu-Jahnes, A. Gupta, K. G. Stawiasz and S. J. Wind in *IBM Journal of Research and Development*, Vol. 39, No. 6, pages 655–668; November 1995.

EXPERIMENTS PROBE THE WAVEFUNCTION OF ELECTRON PAIRS IN HIGH-T_C SUPERCONDUCTORS. Barbara Levi in *Physics Today*, Vol. 49, No. 1, pages 19–22; January 1996.

ELECTRON PAIRS AND WAVES: TACKLING THE PUZZLE OF HIGH-TEMPERATURE SUPERCONDUCTIVITY. Ivars Peterson in *Science News*, Vol. 149, No. 10, pages 156–157; March 9, 1996.

QUESADA/BURKE STUDIOS

The Mystery of Lambic Beer

An ancient brewing technique produces a beverage so complex that it is still yielding its secrets to organic chemists

by Jacques De Keersmaecker

Over the past 2,000 years, the region that is now Belgium has been one of the most heavily trafficked crossroads of Europe. In 57 B.C. the territory, occupied by Celts, was invaded by the Romans and later by Normans. Throughout the centuries, a veritable parade of rulers came and went, including (after the Romans) the Merovingian Franks, the Germans, the Dutch, the Burgundians, the Austrians, the Spanish, the French and the Dutch again. From each of them came the threads, patterns and cloth that created the rich tapestry of Belgian culture, with its diverse contributions to art, cuisine and, not incidentally, brewing.

Some historians suspect that one of these subjugators, possibly the Romans, brought to the region a brewing process that may already have been ancient when it began flourishing around Brussels many centuries ago. During this process, brewers exposed their concoction to the air, causing it to be seeded (or, more accurately, inoculated) by whatever wild, wind-borne yeasts happened to drift in. Only within a roughly 500-square-kilometer area around Brussels and in the Payottenland, a valley of the Senne River on the west side of the city, did the right mix of airborne spores ensure that this spontaneous fermentation occurred consistently. Like Brussels itself, the brewing style thrived amid fields, orchards and woodlands and was nurtured by them.

Centuries later, what is regarded as the oldest surviving commercial brewing style still produces many of the same beers in essentially the same way. This diverse family of beers, known as lambic, includes tart but smooth members, older specimens that are full of complex character, and sweeter fruit- or cane-sugar-flavored varieties. In the better examples is an earthiness—a faint, slightly musky remnant of their wild origin that often comes as a surprise even to jaded beer connoisseurs.

It is likely that all beer was once spon-

HIGHLY AROMATIC LAMBICS are always served, in their native Belgium, in glasses designed to convey their aromas. Fruit lambics, such as cherry, peach, raspberry and plum, are usually poured into snifters or flutes. More traditional gueuze and faro are often served in tumblers.

taneously fermented, like lambic, with wild yeasts only. A brew called *sikaru,* for instance, was produced 5,000 years ago by Sumerians in Mesopotamia. Instead of hops, of which they had no knowledge, Sumerians flavored their brew with spices such as cinnamon. Over the centuries, most of the world's brewers began using techniques that minimized and eventually eliminated the effects of wild yeasts, culminating in the 19th century with the use of scientifically isolated yeast cultures. Fermentation became more efficient and predictable. Through all this, lambic has endured, a throwback to brewing's splendidly eccentric roots.

In 16th-century Belgium, lambic beer soon became a staple of social life. In many paintings of the time by Pieter Brueghel the Elder and other Flemish artists, villagers can be seen quaffing great jugs of lambic, then known as yellow beer. Even today a lambic brewed locally in small quantities is apt to be served at a festival in a Belgian village.

Nowadays larger Belgian brewers, too, produce and even export lambic, capitalizing on popular tastes in some places for hearty, unusual beers and traditional foods. Although lambic beers are still most common in Belgium and parts of the Netherlands and France, they are becoming easier to find elsewhere in Europe and in parts of North America. Given the time-consuming, idiosyncratic nature of lambic brewing, however, the supply is comparatively limited. The output of all lambic breweries adds up to about 370,000 hectoliters a year. In comparison, Boston Beer Company, the 10th largest U.S. brewer and the producer of the Samuel Adams line, brews about 1.1 million hectoliters a year.

Scientific, as well as gustatory, attention has been devoted to lambic lately. Yeast, it turns out, profoundly affects the character of the brew it goes into, and lambic's wild yeasts initiate a fermentation that is complicated, quirky and unfocused. Indeed, lambic arises from a series of stages in which bacterial activity—anathema to other brewing processes—follows yeast fermentation in a sprawling chain of events that produces the relatively large amounts of various sugars and the numerous fragrant compounds, called esters, that give lambic its complex, fruity tastes. Research that began in the mid-1970s at the University of Leuven has finally elucidated much of the chemistry of this centuries-old fermentation.

A Beer Is Born

So little is known about the origins of lambic that there are three different versions of how the word itself came to be. It might have come from any one of four Belgian villages: Lembeek, Borcht-Lombeek, Onze-Lieve-Vrouw-Lombeek or Sint-Katelijne-Lombeek. Another possibility is the Spanish word *lambicado,* which means "carefully prepared." The creation of lambic has also been attributed to Duke Jean IV of Brabant, who in 1428 supposedly tired of the same old brew and hit on the idea of macerating and boiling barley and hops in a still, then known as an *alambic.* The experiment was a success, and the resulting beer has ever since been known as lambic, according to this version.

Old, unblended lambic, close to what was consumed centuries ago, is now easily found only around Brussels and in

the Payottenland. Tart and barely carbonated, it tastes more like fine sherry than beer. Much more common are gueuze (pronounced "gerz"), faro and the various lambics sweetened and flavored with fruit.

Gueuze, like champagne, is the product of a secondary fermentation process. It takes place when young and old lambic are mixed in a bottle. Gueuze was apparently first produced commercially early in the 19th century, to make a more bubbly, beerlike beverage suitable for export. Documents have been found attesting to the export, in 1844, of gueuze to Constantinople and to Rio de Janeiro.

Faro is a blend of lambic and mars, a weakly alcoholic and pale liquid obtained by rinsing the grist of a lambic brew. It is generally sweetened with brown crystallized cane sugar. The name comes from the soldiers of the 16th-century emperor Charles V, who called the product "gold liquor" or "barley liquor"—*farro* in Spanish. Fruit lambics include the traditional cherry (known as *kriek*, the Flemish word for "cherry") and raspberry (or *framboise*, from the French). Other fruits have also been used, with varying degrees of success. They include peaches, grapes, black currants, plums and pineapples.

Essentially all beers—with the possible exception of lambic—are either ales or lagers. All share certain basic kinds of raw materials, such as malt and hops. Malt is barley grain that has been steeped in water, has germinated and has then been dried in a kiln. Malting produces in the grain the enzymes necessary to transform starch into sugar during brewing. The process can be varied to produce certain desired characteristics; lambic malt, for example, is pale and very rich in enzymes. Hops, which are derived from a spicelike plant, are available in dozens of varieties and several different forms. The most popular form for lambic brewing are the dried petals of the hop flowers, also known as cones.

In the Beginning, the Wort

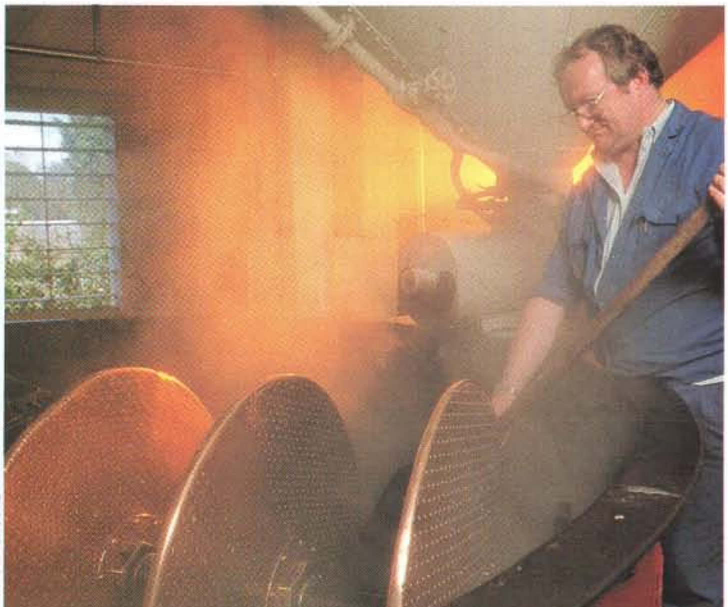
All beer starts off in the same way. Malted grain is boiled—that is, brewed—in a cooker with hops and perhaps with some unmalted grains as well. The malt, grain and water mixture is called a wort. It is boiled with the hops for an hour or more, laying the foundation for the two basic taste elements of the finished beer: fruitiness and sweetness, from the malt and grain, and dryness and bitterness, from the hops.

After brewing comes fermentation, in which colonies of yeast, a single-cell living organism, break down sugars from the malt and grain into ethanol, carbon dioxide and other by-products. Fermentation occurs in the absence of oxygen. Yeast exists in many different strains, each of which gives a characteristic flavor. Indeed, one of the many distinctions between ale and lager is the type of yeast used to inoculate the wort; depending on which kind is used, the yeast settles near the top or the bottom of the vessel after fermentation. Brewers jealously guard their yeast strains because of the strains' important role in establishing the beer's flavor.

Such factors apply to all beers, but it is the specifics—the kinds and proportions of water, malt, grains, hops and yeast, the duration of brewing and fermenting, the maturation and, perhaps, blending—that give beer its diversity. Lambic brewers, in particular, are a possessive, secretive lot, guarding recipes that have been refined over decades, if not centuries.

A few important distinctions separate lambic from other beers. For example, lambic is always brewed with a relatively high proportion of unmalted wheat, usually around 35 percent. (German beers, in contrast, never use unmalted grain.) This grain is relatively high in protein and starch, which are not usu-

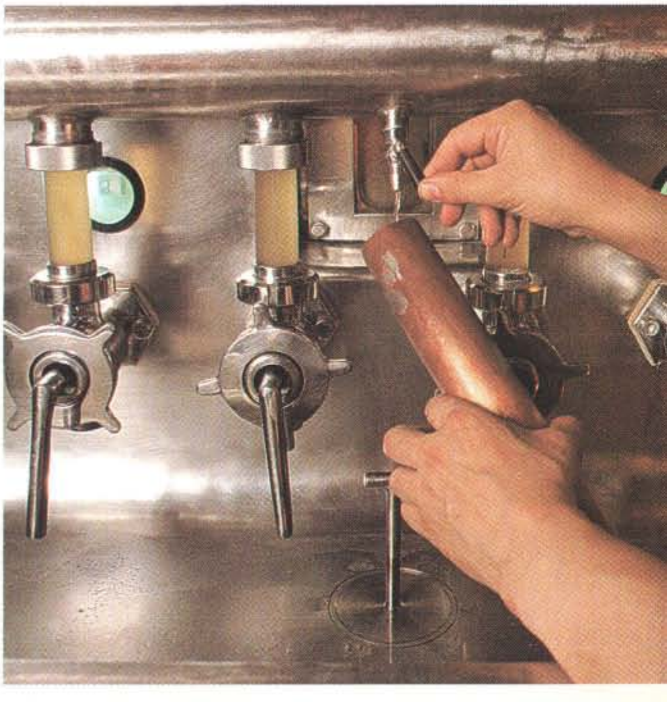
How Lambic Beer Is Made



BERT VAN DEN BROUKE

Brewing begins when a mixture of barley malt and wheat is mashed with warm water in a mash tun (*above*). The round metal plates stir the infusion, helping to extract the flavor and enzymes of the grains. The brewer adds hot water to this mixture, called a wort, then decants the liquid into a cooker. This liquid is warmed in the cooker to activate the enzymes.

After the mashing and warming, the grain-infused liquid is filtered. Valves (*below*) control the rate of flow through the filters, enabling the brewer to maintain the desired clarity.



ally desirable in brewing. High-protein content leaves lambics slightly hazy; on the bright side, it also prolongs a foamy head.

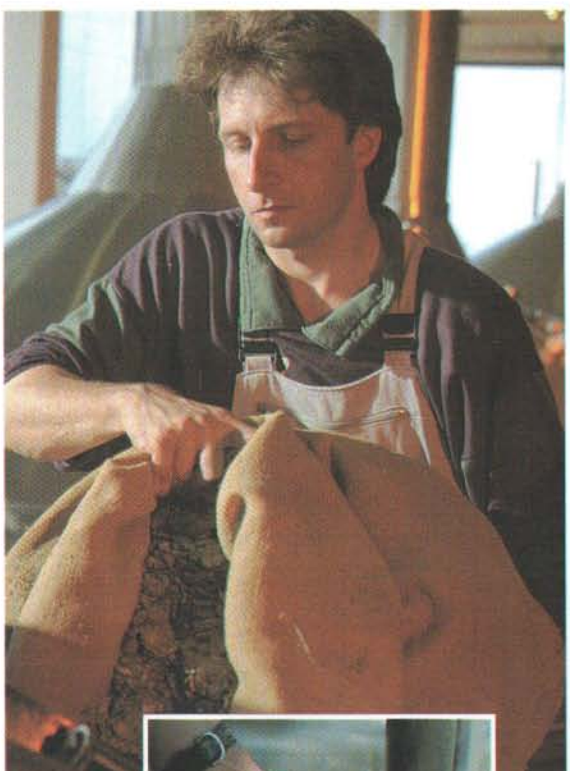
The problem with the wheat starches is that early in the fermentation they lead to high levels of dextrin, a kind of carbohydrate that cannot be broken down by the yeast. Bacterial activity during lambic fermentation does eventually reduce and eliminate dextrins. In moderate amounts, they lend a certain smoothness to a lambic that has fermented for less than a year; by three years, all of them are gone. Unmalted wheat is also relatively low in the enzymes needed for successful brewing—which is why lambic brewers rely on the highly enzymatic, pale barley malt.

Lambic brewing is equally distinctive in its use of hops. In this old brewing style, hops not only flavor but also preserve: their resins contain compounds that inhibit the proliferation of the bacteria responsible for spoilage. Preservation demands relatively great quantities of hops. Ordinarily, such amounts would make the beer too bitter, so lambic brewers use only aged hops at least three years old.

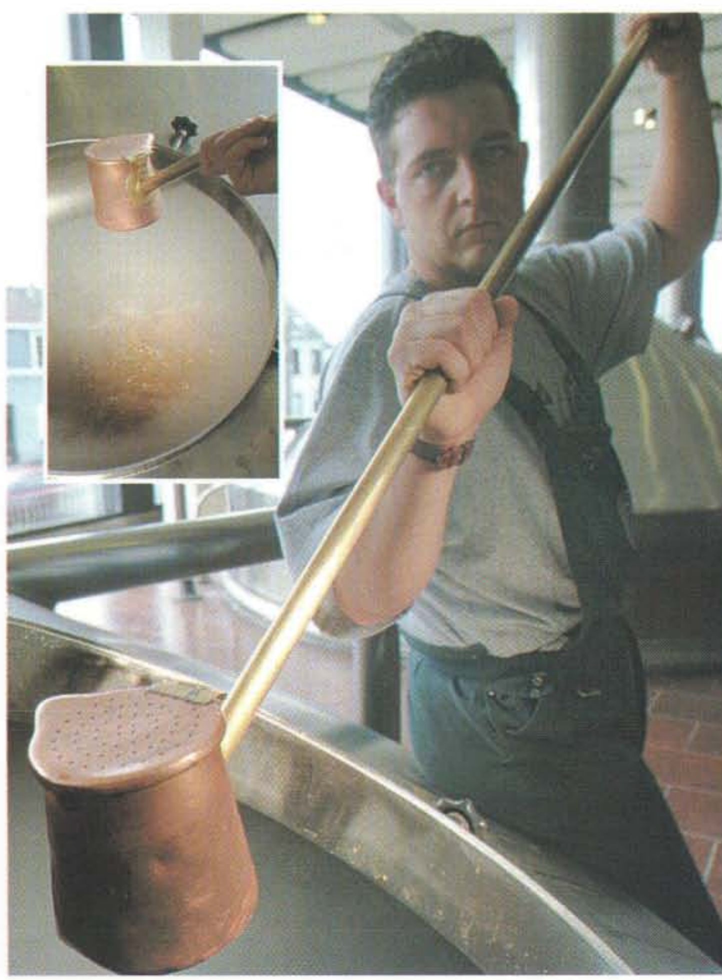
Aging causes the alpha acids, the main source of bitterness, to oxidize and become less bitter. Aging also provokes the oxidation of the hop resins and develops unpleasant haylike and cheesy aromas, but long boiling of the wort eliminates these odors. The traditional favorite is a variety called Coigneau, cultivated in the Asse-Alost area in Belgium. Belgian hops are scarce, however, so hops from the Kent region of England are often substituted.

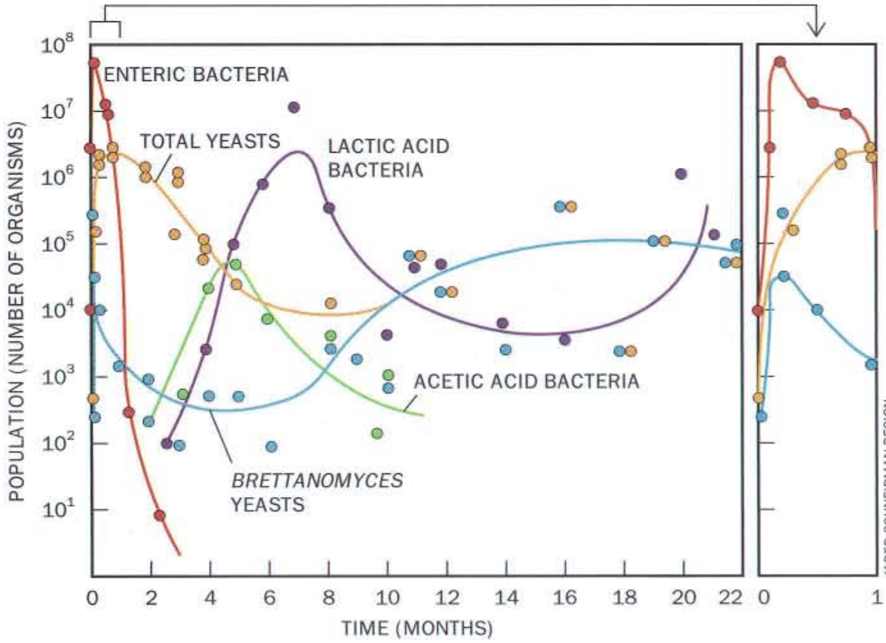
The traditional brewing method, called turbid mashing, has hardly changed since it was described in a book by the brewer G. Lacambre in 1851. In a vessel called a mash tun, the brewer mashes a mixture of about 40 percent wheat and 60 percent barley malt with water at 40 to 45 degrees Celsius. Added boiling water then brings the temperature up to 62 degrees C, after which the liquid is decanted into a cooker. The brewer then pours more boiling water into the mash tun to reach a temperature of 72 degrees C and mashes the grist a second time. After decantation, the liquid phase is also poured into the cooker. The contents of the cooker—liquid produced by the two successive mashings—are then boiled for 20 minutes and poured again into the mash tun for filtration.

The wort resulting from this filtration goes on eventually to become lambic. The grist left in the mash tun is rinsed with hot water, to produce a second



After filtration, the liquid is boiled in the cooker for four or five hours, during which time the brewer adds hops (*top left*). For lambic, brewers prefer the dried petals of hop flowers (*bottom left*), often from England. They use large quantities of hops because the plant preserves as well as flavors the finished lambic. Mellowed, aged hops prevent the beer from becoming too bitter. During the boiling, samples are taken (*below*) to measure the concentration of sugars.





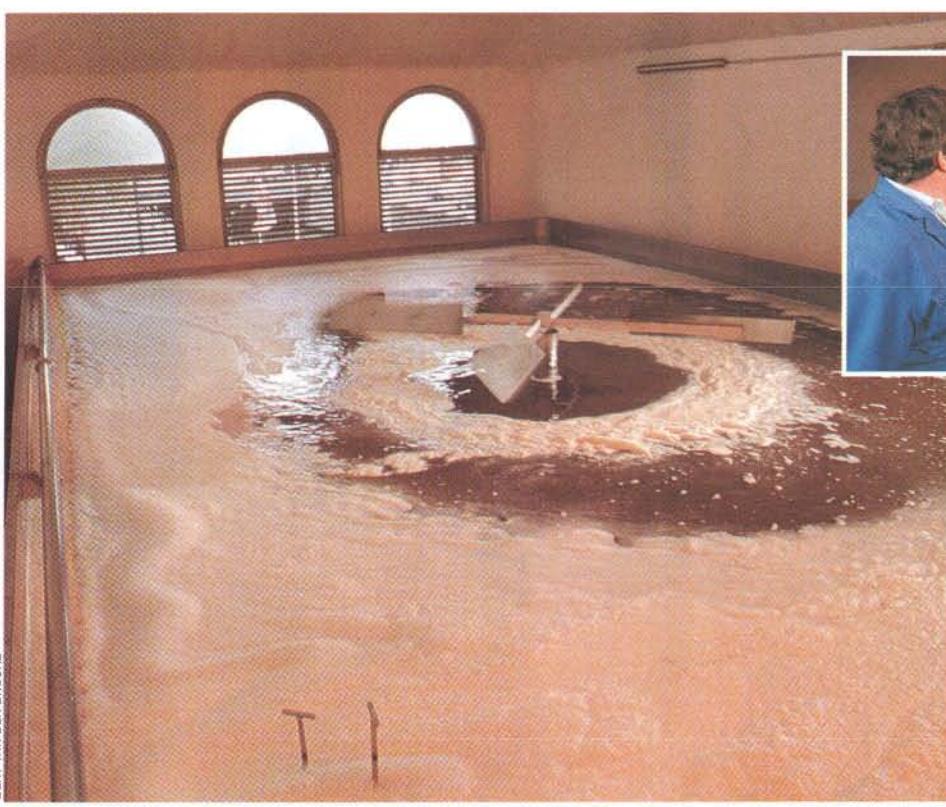
JARED SCHNEIDMAN DESIGN

LAMBIC FERMENTATION encompasses the rise and fall of many different populations of yeast and bacteria in four basic stages. In the first, enteric bacteria and wild yeasts predominate and break down glucose into ethanol, carbon dioxide and acids. Then various yeasts create additional ethanol. In stage three, lactic and acetic bacteria make more of these acids. Finally, *Brettanomyces*, a yeast genus, creates the many esters that make the beer uniquely aromatic.

wort to make mars. The other (lambic) wort goes back into the cooker, flower-like hops are added, and the mixture is boiled for four or five hours. A similar procedure produces mars. After the worts have boiled sufficiently, the brewer filters them to remove the hops.

At this point, the brewing process is complete, and fermentation—the seething riot of chemical and bacterial reactions that actually creates the lambic—begins. It starts with lambic's signature event, unique in all of beer making: the pumping of the hot wort into an open, shallow cooling vessel (also called a tun) in the attic of the brewery. The brewer throws open vented windows, turns on fans and leaves the liquid overnight to cool and be inoculated by the yeasts and other microbial flora of the surrounding air. This exposure to the air is called pitching.

The local conditions are of fundamental importance in pitching. Not only does spontaneous fermentation of wort take place consistently only in a small area around Brussels, but it does so only from October until about April, when outside temperatures remain under 15 degrees C. Some seemingly minute conditions that could affect the balance of microbial flora and the growth rate of the microorganisms would also affect the fermentation sequence and, therefore, the final product. Just how minute some of these conditions may be is a matter of conjecture. Stories are told of one lambic brewer who was faced with a rickety roof that, he was convinced, harbored some critical colonies of microorganisms. So



When the long boil is done, the liquid is pumped into a cooling tun in the brewery's attic (*left*). This is the unique, signature event in the creation of lambic, when local microflora, including wild yeasts, inoculate the brew. Fans circulate the air while the liquid's temperature drops to about 15 degrees Celsius (*above*).

BERT VAN DEN BROUKE

rather than replacing the old roof, he had a new one built over the top of the old one.

After the inoculated wort has cooled, workers pour it into wooden casks for fermentation and maturation. They always use casks of oak or chestnut that had previously held wine, because the wood of new casks would impart an inappropriate tannic taste. Before filling the casks with the wort, the workers scrape and clean the vessels and burn a sulfur wick inside each one to eliminate undesirable molds, which would impart a rotten taste. Even after such cleaning and fumigating, living yeast cells and spores remain within the wood fibers and help to start the fermentation.

Fermentation, which is fairly complex even with the specially cultured yeasts used to make ales and lagers, is much more so with the wild yeasts and bacteria that give rise to lambic. The bacterial activity presents no danger to human health, because the low pH and the presence of hop resins and alcohol combine to eliminate any harmful microorganisms. This, by the way, helps to explain why beer saved countless people from the epidemics of the Middle Ages.

Hubert Verachtert and his colleagues at the University of Leuven have been studying the organic chemistry of lambic fermentation since the 1970s. At last count they had identified 100 different kinds of yeast colonies, 27 colonies of acetic bacteria and 38 colonies of lactic bacteria in a single type of lambic. The difference between the two kinds of bacteria depends on the kind of acid they produce: one creates acetic acid under aerobic conditions; the other makes lactic acid anaerobically.

Yeast Feast

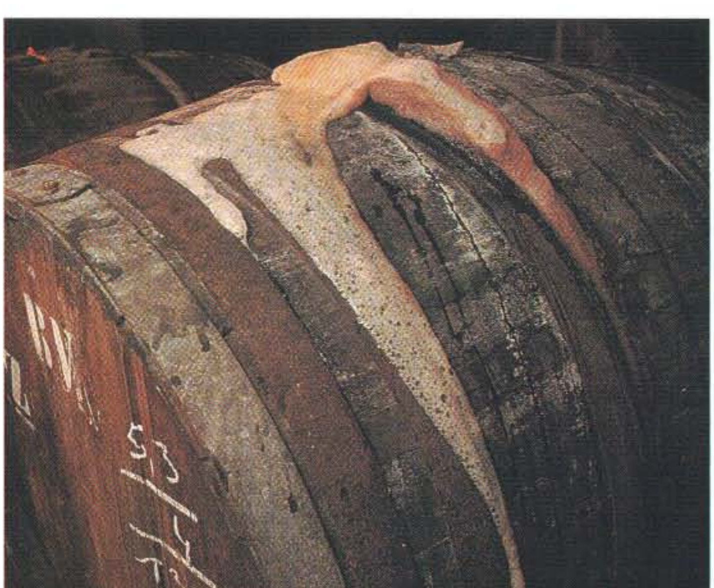
Basically, lambic fermentation takes place in four overlapping stages. In the first, wild yeasts (*Kloeckera apiculata*) and enteric bacteria are the dominant agents. The latter produce a sweet, fruity or vegetablelike aroma. *Kloeckera* yeasts have little influence on taste but foster the breaking down, by enzymes, of proteins, which improves the clarity of the beer at low temperatures. In the second stage, *Saccharomyces* yeasts produce all the ethanol that will be present in the finished product and also create aromatic esters that give the beer flavors similar to those of ale. In stage three, lactic bacteria do what they do best: they make lactic acid, imparting to the beer its characteristic sourness. Finally, *Brettanomyces* yeasts give rise to numerous additional esters, bestowing on lambic its signature fruity, winey taste and aroma.

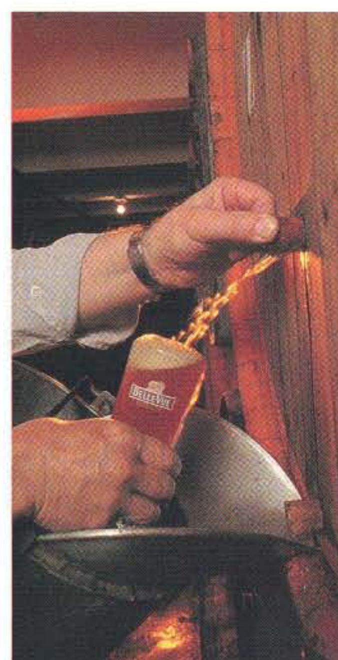
The first stage starts after three to seven days with the proliferation of the enteric bacteria and wild yeasts. Enteric bacteria multiply very quickly and give rise to a so-called mixed-acid fermentation, which means that they break down glucose not only into ethanol and carbon dioxide but also into lactic acid and acetic acid. But because they cannot survive in acetic acid or ethanol, the enteric bacteria population diminishes and then disappears after 30 or 40 days. The *K. apiculata* grows by breaking down glucose but not more complex sugars, so it disappears quite rapidly when the glucose is gone. Before it does, however, its presence produces a small amount of ethanol and generates almost all the acetic acid to be found in the final product.

After 10 or 15 days, the main, alcohol-producing fermentation begins. At this point, the population of wild yeasts is being eclipsed by the proliferation of a different yeast genus, *Saccharomyces*. This genus, incidentally, is the same one used to inoculate ales, in which case they are *S. cerevisiae*, and lagers, in which case they are *S. carlsbergensis*. (Ales and lagers ferment for only a week.)

In lambic, *Saccharomyces* yeasts assimilate and ferment most of the wort's sugars, breaking them down into alcohol and carbon dioxide. The fermentation manifests itself for a couple of days

After the brew has cooled, usually overnight, it is poured into wooden casks (*left*) that were used previously to ferment wine. (The wood of new casks would impart an unpleasant taste.) Even after they are cleaned and fumigated, the casks contain yeast spores that help to start the fermentation. After about a week, the fermentation is so energetic that foam sputters out of escape valves in the casks (*below*).





The beer matures for one or more years. Then, based on taste testing alone, the brew may be blended with younger or older lambics.

as foam sputters and overflows out of the cask's carbon dioxide escape hole. This stage goes on for about seven months. Vigorous though the fermentation is, it is not what brings about the characteristic flavors of lambic, although even at six months some lambics are ready for use in blending the final product.

Stage three, which overlaps stage two, begins after three or four months. This period is characterized by the proliferation of both lactic and acetic bacteria, which reaches a peak after six to eight months. In fact, another reason why lambic brewing always starts between October and April is so that this peak is later assured by summer temperatures above 20 or 25 degrees C.

The lactic acid bacteria created during this phase are responsible for the sour-vinous character of lambic. They are mostly of the *Pediococcus* genus, which converts sugars to lactic acid. Annoyingly, some strains of *Pediococcus* tend to form slime. Fortunately, it eventually disappears, although a slight haze persists. The haze, called double face, cannot be eliminated by filtration.

Acetic bacteria also have undesirable effects, tending to make the beer acidic ("hard") by production of acetic acid from ethanol. Such problems become serious only in damaged or leaky casks that have allowed in air, which nurtures the aerobic *Acetomonas* bacteria.

After eight months, a new increase of yeast cells signifies the beginning of the fourth, and final, stage of lambic's complex fermentation. These yeast strains, first identified in lambic beer, belong mainly to *Brettanomyces*. They play a critical role in establishing the beer's aromatic profile and, therefore, its flavor. The aromatic profile is determined mainly by the concentrations of the beer's many esters, which are, in turn, fostered by *Brettanomyces*. These yeasts produce an enzyme that promotes the reactions that transform acids and alcohol into esters (and vice versa, a phenomenon known as hydrolysis). The most influential of these esters are ethyl lactate and ethyl acetate, by-products of the lactic and acetic acid of the previous stage.

Another characteristic of *Brettanomyces* yeasts is that, in the presence of air, they form a film on the surface of the beer. Other yeasts, too, participate. Without this film, the beer would oxidize, and the acetic bacteria would run rampant, making vinegar instead of beer. Despite the best efforts of brewers, this can happen from time to time.

Another Fermentation

Even after this complex, yearlong fermentation, the beer is, generally, not yet ready for drinking. As mentioned earlier, the most common unflavored lambic, gueuze, is a blend of young and old lambics, going from smooth (mostly young) to tart and complex (mostly old). The exact blending is decided on the basis of taste testing alone. An additional fermentation occurs when the sugars present in a young lambic, about a year old, encounter the more developed suite of yeasts in an older lambic of, say, two or three years. In the traditional method, the blend is refermented in the bottle.

If older lambic predominates in this refermentation, the *Brettanomyces* thoroughly assimilates the complex sugars, leaving an overattenuated, very dry product. It also becomes very aromatic because of the esterification activity of *Brettanomyces*. These qualities would be less intense in a younger average blend, which would be "softer" and smoother. The taste teams have to be aware of the necessary standards and obtain a completely consistent final product by blending very heterogeneous lambics. Regardless of the final blend, the beer differs from ordinary ales and lagers mainly because it has organic acids and a complex suite of esters.

Currently most products are softer ones, which are filtered and pasteurized. The old blends, refermented in the bottle, need careful handling and storage in a cool, dark cellar for at least a year.

Fruit lambics, too, undergo a secondary fermentation, triggered by the sugar in the fruit. Cherry lambic is perhaps the most traditional. Sour Schaarbeek cherries were used in the past, but they have just about disappeared from the market. So brewers generally use Gorsem cherries, which are larger, juicier and convey a sweeter taste. At cherry-picking time, in July, workers prepare all the casks needed for a complete production year. They put about 80 kilos of fruit, complete with stones, into every 650-liter cask and pour in young lambic of that season. Fermentation then starts again, and the beer is left to ferment for one to two years.

Can Lambic Survive?

Lambic is a living anachronism. The very characteristics that make the beer unique are also liabilities, from a business standpoint. Before the work of Louis Pasteur, lambic was the only beer that could be preserved for months or even years. In a relatively low cost operation, lambic brewers could produce large volumes in the winter and then sell the product all year. Although preservation is no longer a problem, maturing great quantities of the beverage for months or years means that at any given time, much of the lambic brewer's capital is immobilized.

In addition, many modern food laws and regulations require minute control over the entire preparation process. By its very nature, however, spontaneous fermentation is not controllable to any real extent.

Lambic's future rests with adventurous beer lovers and that small but enthusiastic segment of the population that goes out of its way to sample traditional ethnic foods. Lately this group seems to be expanding as more people pass up processed foods in favor of the old staples: fine cheeses, hearty breads, wines, abbey beers and real ales. Who knows? If the trend continues, lambic may be around for another 500 years.

The Author

JACQUES DE KEERSMAECKER is managing director of the Belle-Vue Brewery in Brussels, an affiliate of the Interbrew-Labatt group. A connoisseur and gustatory adventurer, he was born in a lambic brewery and studied at the Brussels Institute's Meurice Brewing Engineers' School. He was formerly director of the Lamot Brewery in Mechelen, Belgium.

Further Reading

LES MÉMOIRES DE JEF LAMBIC. La Technique Belge Editions, Brussels, undated.

LAMBIC. Jean-Xavier Guinard. Classic Beer Style Series. Brewers Publications, Boulder, Colo., 1990.

MICHAEL JACKSON'S BEER COMPANION. Michael Jackson. Running Press, 1993.

LAMBIC: BELGIUM'S UNIQUE TREASURE. Martin Lodahl in *Brewing Techniques*, Vol. 3, No. 4, pages 34–46; July–August 1995.

KEN MARTEN *Earthtrust*

Ring Bubbles of Dolphins

KEN MARTEN, SUCHI PSARAKOS AND DON J. WHITE *Earthtrust*

A number of bottlenose dolphins in Hawaii can create shimmering, stable rings and helices of air as part of play

by Ken Marten, Karim Shariff, Suchi Psarakos and Don J. White

Below the towering cliffs of Makapuu Beach on the island of Oahu, Hawaii, is a unique laboratory dedicated to the study of dolphins. Project Delphis, run by the nonprofit conservation organization Earthtrust, in cooperation with Sea Life Park Hawaii, conducts research ranging from investigating dolphin self-awareness to exploring the animals' intelligence using an underwater computer touch screen. The scientists in the lab do not use food as a reward, so all the behavior observed is of the dolphins' own volition.

One of the most fascinating activities we have seen in our research involves no high-tech human toys at all. Instead the dolphins fashion their own entertainment by swirling the water with their fins and blowing bubbles into the resulting vortices to produce rings and helices of air. Furthermore, the physics behind the air rings turns out to be quite interesting. Few people doubt that dolphins are highly intelligent animals, but these observations demonstrate just how imaginative they can be.

As air breathers, dolphins blow bubbles whenever they exhale underwater. Dolphin behaviorists have noted that when dolphins are excited, surprised or curious, they will sometimes expel air from their blowholes, generating large, amorphous bubbles that rise quickly to the surface. The animals occasionally emit streams of small bubbles when they make sounds; the bubbles might add another component, detectable by sight or sonar, to the vocal message.

KAIKO'O (*left*), an adult bottlenose dolphin at Sea Life Park Hawaii, generates a rising ring of air to play with. The laboratory at the park (*above, right*) offers a window into the world of dolphins.

Dolphins can also create more exotic types of bubbles for less prosaic reasons. In recent years, researchers at several oceanariums around the world have reported that a variety of marine mammals can blow smooth, stable rings of air that linger in the water for several seconds. Because of the intricate techniques and practice required to form such rings, as well as the helices we have seen, these bubbles are clearly not a spontaneous response to alarm or a standard part of communication. Wolfgang Gewalt of the Duisburg Zoological Gardens in Germany observed untrained Amazon river dolphins (*Inia geoffrensis*) producing bubbles in unusual and playful ways. The animals emitted air from their mouths to yield necklaces of bubbles, which they would pass through or bite.

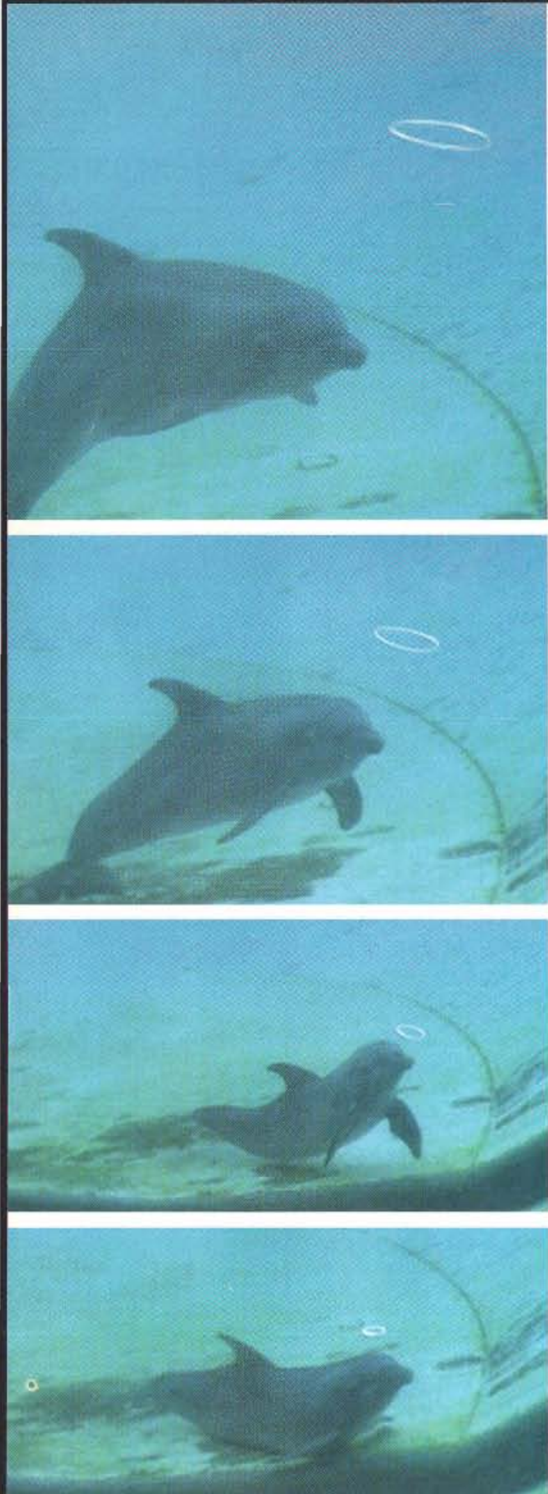
Elsewhere, at Marine World Africa USA in Vallejo, Calif., Diana Reiss and Jan Ostman-Lind noticed bottlenose dolphins (*Tursiops truncatus*) at the aquarium playing with rings similar to the ones we have seen. Kenneth S. Norris of the University of California at Santa Cruz described beluga whales (*Delphinapterus leucas*) at the Vancouver Aquarium that expelled bubbles from their blowholes and then sucked them into their mouths as part of playful behavior. Some people have even witnessed air rings in the wild. The behavioral biologist Karen Pryor observed male Pacific spotted dolphins (*Stenella attenuata*) blowing rings during social encounters; Denise Herzing of Florida Atlantic University has viewed similar displays in free-ranging Atlantic spotted dolphins (*Stenella frontalis*). And the marine mammal photographer Flip Nicklin has seen beluga whales living in Lancaster Sound in Baffin Bay, Canada, that released rings of air as they repeatedly clapped their jaws together, possibly in a display of aggression.

Hawaiian Ring Culture

During the past five years at Sea Life Park Hawaii, we have studied 17 bottlenose dolphins; nine of them, ranging in age from 1.5 to 30 years old, generated air rings. Based on our observations at other oceanariums and consultations with colleagues at various sites, we believe ring blowing is more common at Sea Life Park Hawaii than at other aquariums; the dolphins here appear to have created a "ring culture" in which novice dolphins learn to make rings in the presence of experts that, in a sense, pass down the tradition.

Ring making is a leisurely pastime, so the animals generate rings only when they want to—not on command or for a reward of food. Furthermore, ring making does not seem to be associated with functional behaviors such as eating or sexual activity. Because ring blowing is unpredictable, we have had some difficulty documenting it. But over time we have captured most of the dolphins at play in photographs and on videotape. (Additional photographs can be seen on the World Wide Web at http://earthtrust.org, and a Quicktime movie can be found at http://www.sciam.com/) Unfortunately, neither of these media does the dolphins justice. But the archived images do reveal important information about the physics of the rings.

From what we have seen, the dolphins employ three basic techniques to form the rings. In the simplest method (also



HALO OF AIR becomes thinner and expands in radius as it is carried to the surface by the vortex flow (*arrows*) through the center of the ring. The vortex also serves to stabilize the ring, preventing it from breaking up into smaller bubbles. Here Kaiko'o produces a ring and follows it up through the water (*video stills, bottom to top*).

used by human divers), dolphins puff out bubbles from their blowholes; these bubbles become halos of air that expand in radius while decreasing in thickness as they rise to the surface [*see illustration above*]. One of the dolphins we watched, the adult male Kaiko'o, could emit two rings in succession and then fuse them into a single, large ring.

The physics behind this type of ring is relatively straightforward: any spherical bubble bigger than about two centimeters in diameter will quickly become a ring because of the difference in water pressure above and below the bubble. Water pressure increases with depth, so the bottom of the bubble experiences a higher pressure than the top does. The pressure from below overcomes the surface tension of the sphere, punching a hole in the center to create a doughnut shape.

As water rushes through this hole, a vortex forms around the bubble. Any vortex ring travels in the same direction as the flow through its center; in the case of these simple air rings, the vortex flow, in combination with the air's natural buoyancy, propels the bubbles toward the surface. Although the process of making air-filled rings is fairly simple, dolphins cannot blow stable ones without some practice, suggesting that additional factors—such as the viscosity of the water and how the air is ejected from the blowhole—need to be taken into account.

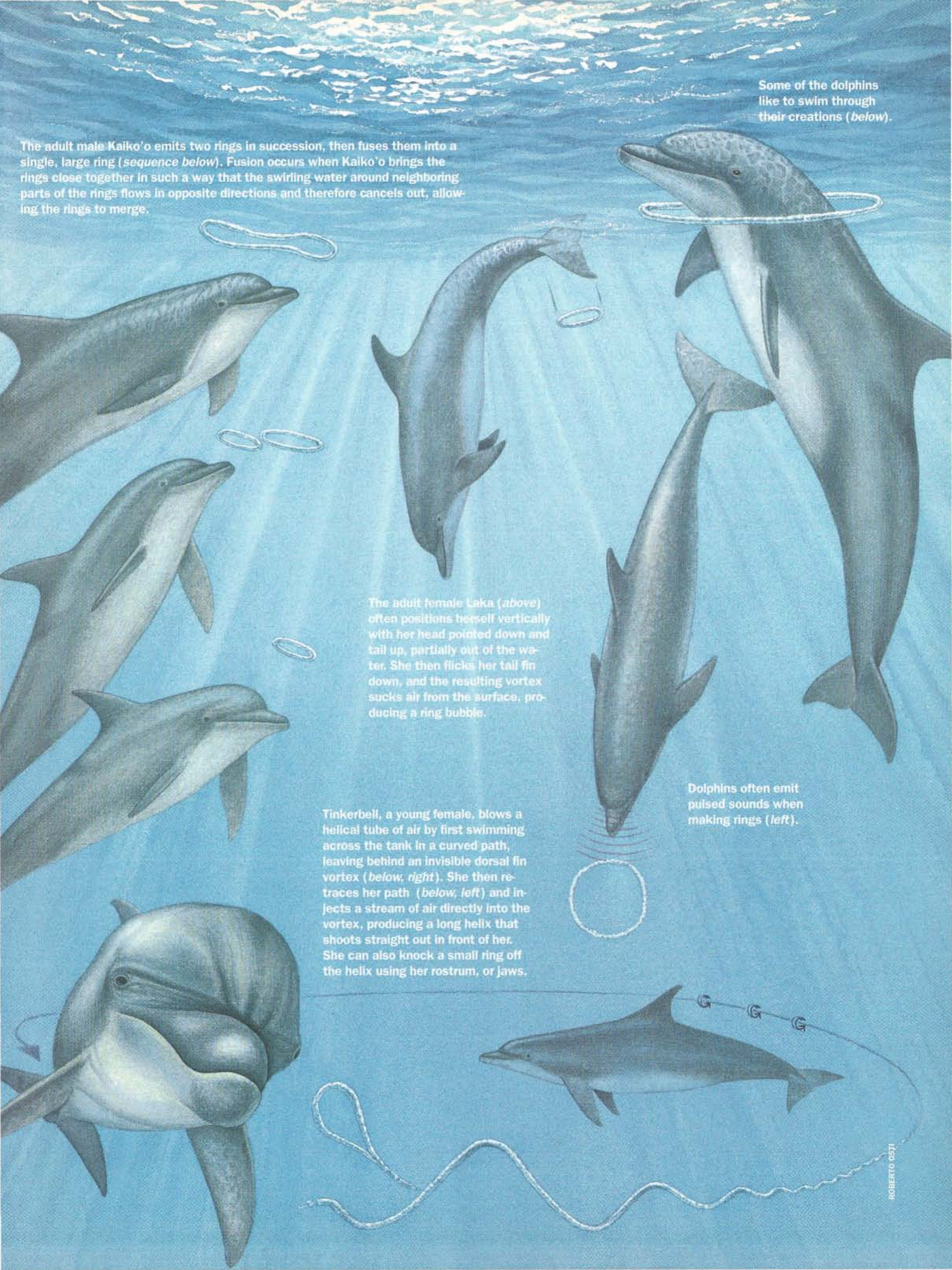
In a more elaborate approach, the dolphins fabricate rings that travel horizontally and sometimes even downward in the water. For instance, a dolphin might swim forward rapidly on its side so that its normally horizontal flukes (that is, its tail fin) are vertical. By thrusting its flukes vigorously to one side, the animal generates an invisible, ring-shaped vortex that travels horizontally and slightly downward. After quickly turning around, the dolphin finds the vortex and injects a bubble into it from the blowhole. (The dolphin often produces an audible series of clicks before the release of air, suggesting that sonar may be employed to locate the invisible vortex.) The pressure inside a vortex is lowest in the center, or "eye," of the swirl; when the dolphin exhales into the vortex, the air migrates to the region of lowest pressure and is drawn out along the core of the ring-shaped vortex [*see illustrations at top of pages 68 and 69*].

The resulting ring can be up to 60 centimeters (two feet) across and just over a centimeter thick, traveling horizontally in the water. Once again, the movement of the ring reflects the direction of the flow through its center; in the case of the vortex created by the dolphin tail fin, this flow is horizontal and sometimes even downward. With a sideways toss of its rostrum, or jaws, the dolphin can pull a small ring off the larger one and then steer it through the water.

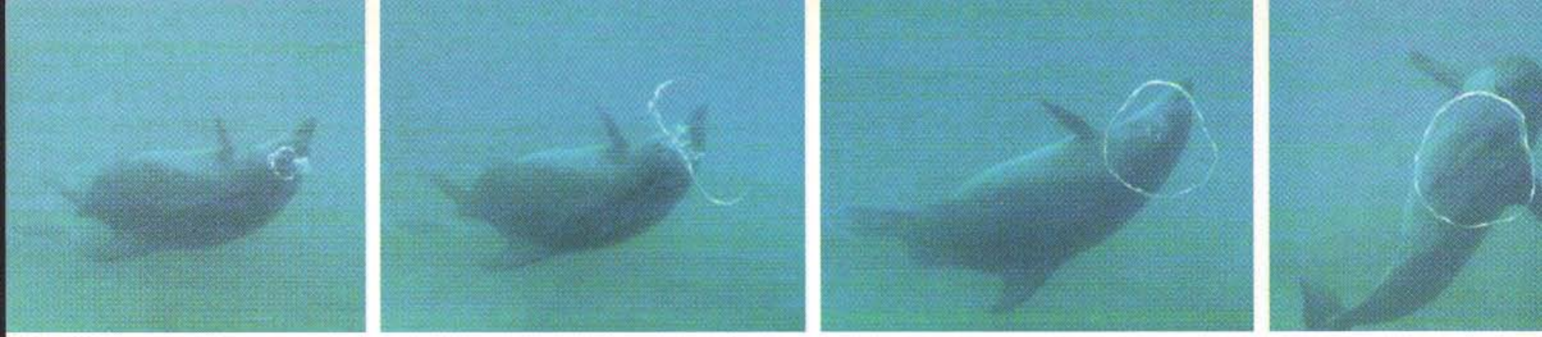
Making a vortex with flukes and planting a bubble in it can be done in a variety of ways—the adult female Laka often positions herself vertically in the water with her head pointed down and tail up. Laka then flips her tail fin down to stir up a vortex. She fills the vortices not only with air from her blowhole but also with air from her mouth. Occasionally, Laka will capture air from above the water with a downward thrust of her flukes. In addition, we have watched Laka release from her mouth small bubbles that pass along her body; when the bubbles reach her flukes, she flips them into a ring. And she can even augment a ring by injecting more air into it.

Experimenting with Bubbles

The third type of air-filled vortex dramatically reveals the dolphins' capacity for experimentation. On a few occasions, we watched the young female dolphin Tinkerbell, Laka's daughter, construct long helices of air, using the most complicated technique we have seen. These more complex structures no doubt result from considerable refinement through trial and error. Tinkerbell has developed two very different methods for making helices. In one approach, she releases a group of small bubbles while swimming in a curved path near the wall of the tank. She then turns quickly, and as the dorsal fin on her back brushes past the bubbles, the vortex formed by the fin brings the bubbles together and coils them into a helix three



Some of the dolphins like to swim through their creations (*below*).
The adult male Kaiko'o emits two rings in succession, then fuses them into a single, large ring (*sequence below*). Fusion occurs when Kaiko'o brings the rings close together in such a way that the swirling water around neighboring parts of the rings flows in opposite directions and therefore cancels out, allowing the rings to merge.
The adult female Laka (*above*) often positions herself vertically with her head pointed down and tail up, partially out of the water. She then flicks her tail fin down, and the resulting vortex sucks air from the surface, producing a ring bubble.
Dolphins often emit pulsed sounds when making rings (*left*).
Tinkerbell, a young female, blows a helical tube of air by first swimming across the tank in a curved path, leaving behind an invisible dorsal fin vortex (*below, right*). She then retraces her path (*below, left*) and injects a stream of air directly into the vortex, producing a long helix that shoots straight out in front of her. She can also knock a small ring off the helix using her rostrum, or jaws.
ROBERTO OSTI



to five meters (10 to 15 feet) long [*see illustrations at bottom of these two pages*]. We have also seen Tinkerbell swim across the tank in a slightly curved path, leaving behind an invisible dorsal fin vortex. She then retraces the path and injects a stream of air into the vortex, producing a long helix that shoots out in front of her.

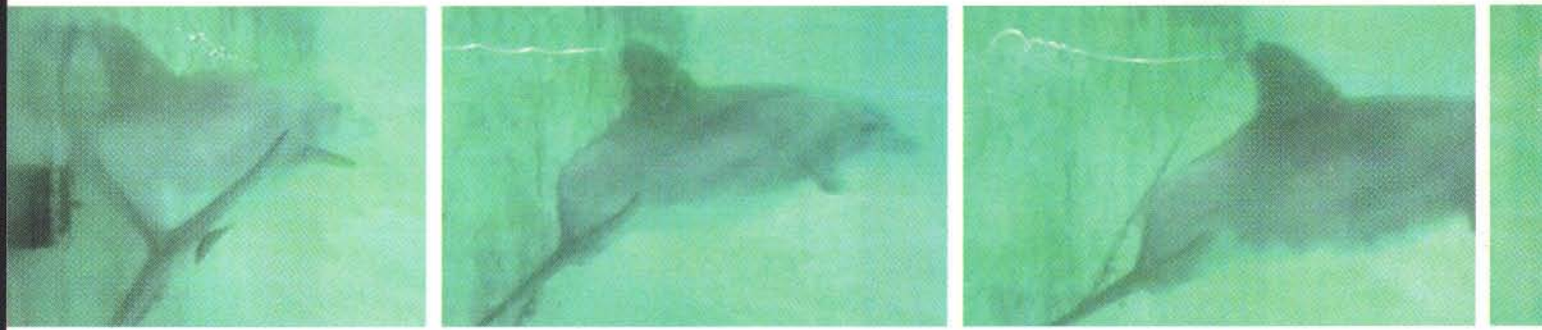
Again, because the pressure in a vortex is lowest in the eye, once the bubbles are inside the vortex, they move toward the center, merge and elongate into a helical tube. Usually a tube of air in water is unstable and breaks up into smaller bubbles. But all the dolphins' rings and helices are shiny and smooth because the variation of pressure inside the vortex (low pressure at the center, building up to higher pressure at the edges) actually works to stabilize the tube by smoothing the ripples that would otherwise break up the large bubble.

Although we did not see any dolphins other than Tinkerbell generating helices, the practice of making rings spread through the population of dolphins, as some of the individuals learned the technique in the presence of their ring-blowing companions. We had the opportunity to watch one young dolphin's rings evolve over a period of two months from unstable, sloppy bubbles that dissipated rapidly to stable, shimmering rings that lingered in the water for several seconds. Older dolphins also needed time to acquire the talent. One adult male, Keola, lived in the research tank for two years with dolphins that did not produce air rings, and during that time we did not see him generate any. But when his younger, ring-blowing sibling Kaiko'o moved into the same tank, Keola watched for long stretches while Kaiko'o blew rings; within a couple of months, Keola began making his own rings, which slowly progressed in quality.

We have noticed that other dolphins also monitor their ring-blowing tank mates intently, suggesting that the exhibition interests the animals or offers a learning opportunity for them. On several occasions we saw the two brothers Keola and Kaiko'o lying side by side on the bottom of the tank, repeatedly blowing large doughnut rings either simultaneously or within a second of each other. We have also seen one female, swimming closely behind another female who was blowing rings, produce her own bursts of small bubbles as she watched.

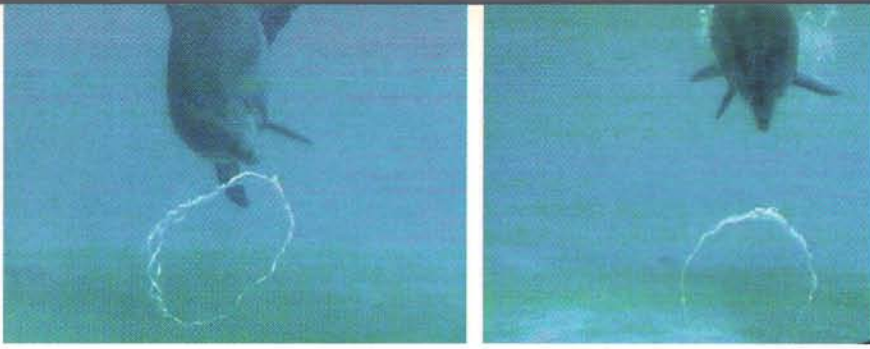
The dolphins have drawn humans into their play as well: one day during a period of intense ring making, Tinkerbell repeatedly blew a ring and then came to the lab window where one of us (Psarakos) was videotaping, as if to include her in the activity. Once, we blew soap bubbles inside the lab in front of the dolphins' window, and within a few minutes one of the dolphins joined in by blowing simple, rising doughnut rings

MARTEN, PSARAKOS AND WHITE *Earthtrust*

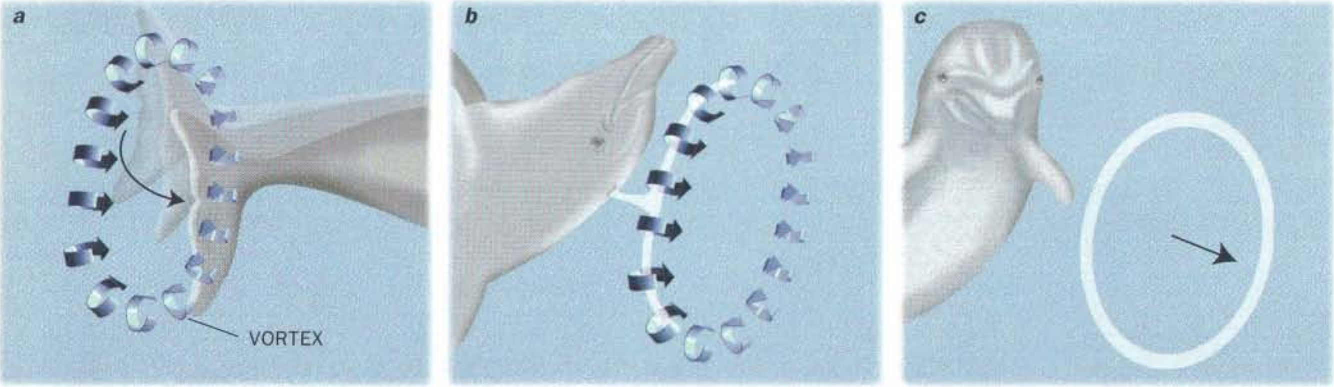


LAURIE GRACE

CORKSCREWS OF AIR are uncommon: the authors have observed only Tinkerbell (*video stills, above*) create them. In one technique, Tinkerbell emits a burst of bubbles while producing a spiral vortex off the dorsal fin (*a*). When the bubbles encounter the vortex (*b*), they are drawn into it, then merge and elongate into a long helix of air (*c*).



HORIZONTALLY MOVING RING forms when a dolphin flicks its tail while swimming on its side. The horizontal and slightly downward motion of the fin creates a vortex traveling in the same direction (*a*). The dolphin then turns around and injects air into the swirling flow (*b*); the air is drawn along the core of the vortex (*c*), forming a ring that moves in the direction of flow through its center. The adult female Laka is pictured (*left*) exhaling into the vortex and examining her creation.



LAURIE GRACE

near the lab window. The real surprise came when the dolphin swam away from the window and made several fluke vortex rings—so different from what we each blew at the window.

Our study of the bottlenose dolphins at Sea Life Park Hawaii continues in the hope of better understanding their behavior. As the only nonprimates that have shown strong indications of self-awareness, these dolphins may teach us about the nature of intelligence through their experimentation and play. But as we consider their remarkable abilities, we are haunted by the knowledge that many cultures, including our own, regard dolphins as expendable. Dolphins continue to be targeted by tuna nets, to become ensnared in expansive drift nets and gill nets, to be canned as mock whale meat and to be shot for crab bait or to be hunted for sport. Earthtrust and its sponsors work to address these issues, but we believe only a basic change in human behavior will make a permanent difference. It is our fervent hope that by providing new views into the dolphin mind, we may yet convince people to stop the indiscriminate slaughter of these fascinating creatures.

The Authors

KEN MARTEN, KARIM SHARIFF, SUCHI PSARAKOS and DON J. WHITE have worked together in the study of dolphin rings for the past two years. Marten, who studies bioacoustics as well as dolphin behavior, is director of research for Project Delphis at Earthtrust. Shariff specializes in fluid dynamics at the National Aeronautics and Space Administration Ames Research Center. Psarakos is the computer scientist and co-director of the wild dolphin research project at Project Delphis. White founded Earthtrust in 1976 to promote the conservation of marine mammals; he is currently spearheading a new campaign in the area of nonintrusive dolphin research and public education. The authors gratefully acknowledge the support of the staff at Sea Life Park Hawaii. Information on Earthtrust is available from earthtrust@aloha.net or 25 Kaneohe Bay Drive, Kailua, HI 96734.

Further Reading

VORTEX RING BUBBLES. T. S. Lundgren and N. N. Mansour in *Journal of Fluid Mechanics*, Vol. 225, pages 177–196; March 1991.

EVIDENCE OF SELF-AWARENESS IN THE BOTTLENOSE DOLPHIN (*TURSIOPS TRUNCATUS*). Ken Marten and Suchi Psarakos in *Self-Awareness in Animals and Humans: Developmental Perspectives*. Edited by S. T. Parker, R. W. Mitchell and M. L. Boccia. Cambridge University Press, 1994.

USING SELF-VIEW TELEVISION TO DISTINGUISH BETWEEN SELF-EXAMINATION AND SOCIAL BEHAVIOR IN THE BOTTLENOSE DOLPHIN (*TURSIOPS TRUNCATUS*), by Ken Marten and Suchi Psarakos, pages 205–224, and related articles on animal consciousness in *Consciousness and Cognition*, Vol. 4, No. 2, pages 205–269; June 1995.

General information on dolphin cognition research and the organization Earthtrust is available on the World Wide Web at http://earthtrust.org

Gaining on Fat

by W. Wayt Gibbs, *staff writer*

PHOTOGRAPHS COURTESY OF LUCY D. LUSTIG-CURTIS

BATTLE AGAINST BIOLOGY often leads to cycles of weight loss and regain. Lucy D. Lustig-Curtis has been through three. She put on weight through childhood, then lost 40 pounds at a diet camp. In college, her weight rose again to 300 pounds, then fell to 185 (*left*). After peaking at 663 pounds in 1994, Lustig-Curtis is now down to about 395 (*right*).

Throughout most of human history, a wide girth has been viewed as a sign of health and prosperity. It seems both ironic and fitting, then, that corpulence now poses a growing threat to the health of many inhabitants of the richest nations. The measure of the hazard in the U.S. is well known: 59 percent of the adult population meets the current definition of clinical obesity, according to a 1995 report by the Institute of Medicine, easily qualifying the disease for epidemic status. Epidemiologists at Harvard University conservatively estimate that treating obesity and the diabetes, heart disease, high blood pressure and gallstones caused by it rang up $45.8 billion in health care costs in 1990, the latest year studied. Indirect costs because of missed work pitched another $23 billion onto the pile. That year, a congressional committee calculated, Americans spent about $33 billion on weight-loss products and services. Yet roughly 300,000 men and women were sent early to their graves by the damaging effects of eating too much and moving too little.

The problem is as frustrating as it is serious. Quick and easy solutions—liquid diets, support groups, acupressure, appetite-suppressing "aroma sticks" and even the best-intentioned attempts to eat less and exercise more—have all failed in well-controlled trials to reduce the weight of more than a small fraction of their obese adherents by at least 10 percent for five years—an achievement shown to increase life expectancy sharply.

The discovery last summer of leptin, a natural hormone that cures gross obesity when injected into mutant mice that lack it, raised hopes of a better quick fix. Those hopes have faded as subsequent studies have found no fat people who share the leptin-related mutations seen in mice. But the identification of leptin is only one of many important advances over the past several years that have opened a new chapter in the understanding of obesity.

Armed with powerful new tools in molecular biology and genetic engineering, scientists are seeking physiological explanations for some of the most puzzling aspects of the fattening of industrial society. Why is obesity on the rise, not just in the U.S. but in nearly all affluent countries? How is it that some individuals remain fat despite constant diets, whereas others eat what they want without gaining a pound? Why is it so hard to lose a significant amount of weight and nearly impossible to keep it off? Perhaps most important, what can be done to slow and eventually reverse this snowballing trend? The traditional notion that obesity is simply the well-deserved consequence of sloth and gluttony has led to unhelpful and sometimes incorrect answers to these questions. Science may at last offer better.

What Makes the World Go Round

Contrary to conventional wisdom, the U.S. is not the fattest nation on earth. Obesity is far more common on Western Samoa and several other Pacific islands. On Nauru, a mere dot of eight square miles once covered to overflowing with seabird guano, the 7,500 island-

As a costly epidemic of obesity spreads through the industrial world, scientists are uncovering the biological roots of this complex disease. The work offers tantalizing hope of new ways to treat, and prevent, the health risks of excess weight

ers have traded that valuable source of phosphate to fertilizer companies in exchange for one of the highest per capita incomes in the world. Many also traded their plows for lounge chairs and their traditional diet of fish and vegetables for Western staples such as canned meats, potato chips and beer. Within the course of a generation, the change has taken its toll on their bodies. By 1987 well over 65 percent of men and 70 percent of women on Nauru were obese, and one third suffered from diabetes.

Many countries, developed and developing, are heading in the same direction at an alarming pace. Changes in diet alone do not explain the trend. Surveys—some of which admittedly are of dubious accuracy—show that the proportion of calories Americans get from fat has dropped about eight points since the 1980s, to 34 percent. Yet the prevalence of obesity has risen by a similar amount in nearly the same period. Britons ate 10 percent fewer calories overall in 1991 than in 1980, according to government estimates, while the number of heavyweights doubled. Polls that show gasoline consumption and hours spent watching television rising about as quickly as the rate of obesity in some countries seem to explain part of the disparity.

Evolutionary biology may provide a deeper explanation, however. In 1962 James V. Neel of the University of Michigan proposed that natural selection pressured our distant ancestors to acquire "thrifty genes," which boosted the ability to store fat from each feast in order to sustain people through the next famine. In today's relative surfeit, Neel reasoned, this adaptation has become a liability. The theory is supported by the Nauruans' plight and also by studies of the Pima Indians, a tribe whose progenitors split into two groups sometime during the Middle Ages. One group settled in southern Arizona; the other moved into the Sierra Madre Mountains in Mexico. By the 1970s most of the Indians in Arizona had been forced out of farming and had switched to an American diet with 40 percent of its calories from fat. They now endure the highest incidence of obesity reported anywhere in the world—far higher than among their white neighbors. About half develop diabetes by age 35.

Eric Ravussin, a researcher with the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), has compared Pimas in Arizona with their distant relatives in Maycoba, Mexico, who still live on subsistence farming and ranching. Although the groups share most of the same genes, Pimas in Maycoba are on average 57 pounds (26 kilograms) lighter and about one inch (2.5 centimeters) shorter. Few have diabetes. Maycobans also eat about half as much fat as their counterparts to the north, and they spend more than 40 hours a week engaged in physical work. The fact that Mexican Pimas remain lean provides strong evidence that the high rate of obesity among American Pimas is the result not of a genetic defect alone but of a genetic susceptibility—exceptionally thrifty genes—turned loose in an environment that offers easy access to high-energy food while requiring little hard labor.

Because all human populations seem to share this genetic susceptibility to varying degrees, "we are going to see a continuing increase in obesity over the next 25 years" as standards of living continue to rise, predicts F. Xavier Pi-Sunyer, director of the obesity research center at St. Luke's-Roosevelt Hospital in New York City. He warns that "some less developed countries are particularly at risk. It is projected that by 2025, more than 20 percent of the population of Mexico will have diabetes."

Studies of Pimas, islanders and migrants "all seem to indicate that among different populations, the prevalence of obesity is largely determined by environmental conditions," Ravussin concludes. A few doctors have proposed changing those conditions by levying a "fat tax" on high-calorie foods or raising insurance rates for those who fail to show up at a gym regularly.

But economic and legal punishments are unlikely to garner much popular support, and no one knows whether they would effectively combat obesity. So most researchers are turning back to factors they think they can control: the genetic and biological variables that make one person gain weight while others in the same circumstances stay lean.

Finding Genes That Fit

Doctors have long known that the tendency to gain weight runs in families—how strongly is still under debate. Numerous analyses of identical

A Shifting Scale

Obesity appears to be rising in most industrial nations, although comparisons are tricky because epidemiologists have never settled on consistent categories for measuring the disorder. Nearly all rely on the body mass index (BMI) [*see formula below*], because this figure is highly correlated with body fat. Still, studies have used a wide range of BMI levels, from below 27 to over 30, to categorize the obese.

The World Health Organization classifies obesity in three levels, with those having BMIs of 30 or higher considered at major risk. Doctors in the U.S. have conventionally used "ideal weight" tables assembled by the Metropolitan Life Insurance Company from actuarial data. Yet recent mortality studies, such as one published last year by Harvard University researchers who examined 115,195 nurses over 16 years, have found that the standard tables underestimate the risks of excess weight—primarily because they fail to account for smokers, who tend to be thin but unhealthy. These newer studies show risks increasing significantly at BMIs of 25 and higher. In 1995 the National Institutes of Health and the American Health Foundation issued new guidelines that define healthy weight as a BMI below 25. According to a recent report by the Institute of Medicine, 59 percent of American adults exceed that threshold.

Calculating Body Mass Index

$$BMI = \frac{w}{h^2}$$

w is weight in kilograms (pounds divided by 2.2)
h is height in meters (inches divided by 39.4)

twins reared apart have shown that genetic factors alone control a large part of one's body mass index, an estimate of body fat commonly used to define obesity [*see box above*]. A few have found weight to be as dependent on genes as height: about 80 percent. But the majority have concluded that genetic influences are only about half that potent.

Investigators at the National Institutes of Health who examined more than 400 twins over a period of 43 years concluded that "cumulative genetic effects explain most of the tracking in obesity over time," including potbellies sprouting in middle age. Interestingly, the researchers also determined that "shared environmental effects were not significant" in influencing the twins' weight gain. That result is bolstered by five studies that compared the body mass indexes of adopted children with their biological and adoptive parents. All found that the family environment—the food in the refrigerator, the frequency of meals, the type of activities the family shares—plays little or no role in determining which children will grow fat. Apparently, only dramatic environmental differences, such as those between the mountains of Mexico and the plains of Arizona, have much effect on the mass of a people.

Just which genes influence our eating, metabolism and physical activity, and how they exert their power, remains a mystery. But geneticists do have some encouraging leads. Five genes that can cause rodents to balloon have now been pinpointed.

Obese, cloned by Jeffrey M. Friedman and others at the Rockefeller University, encodes a blueprint for leptin, a hormone produced by fat cells. Mice with a mutation in this gene produce either no leptin or a malformed version and quickly grow to three times normal weight. *Diabetes*, cloned last December by a team at Millennium Pharmaceuticals in Cambridge, Mass., codes for a receptor protein that responds to leptin by reducing appetite and turning up metabolism. Mice with a bad copy of this gene do not receive the leptin signal, and they, too, get very fat from infancy.

Within the past year scientists at Jackson Laboratory in Bar Harbor, Me., have cloned two other fat genes, named *fat* and *tubby*. Mice with a mutation in either of these genes put on weight gradually—more like humans do. The *fat* gene gets translated into an enzyme that processes insulin, the hormone that signals the body that it has been fed. But the protein produced by the *tubby* gene is unlike any ever seen. Researchers do not yet know why mice with errors in *fat*, *tubby* or *agouti yellow*, a fifth obesity gene discovered several years ago, put on extra ounces.

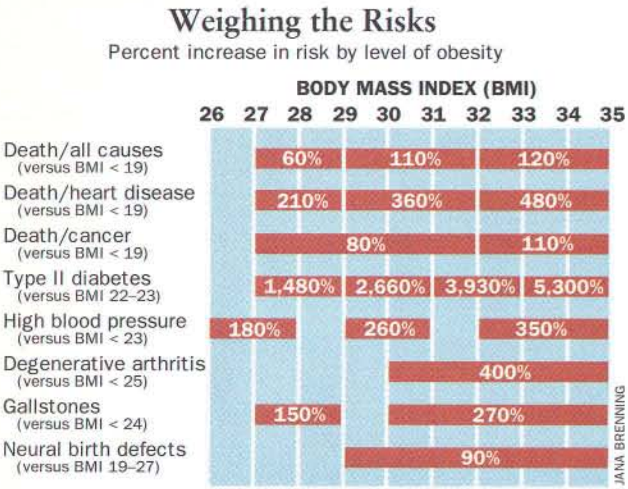
Although geneticists have located versions of all five genes within human DNA, "so far, when we have looked for human mutations on these genes, we haven't found them," reports L. Arthur Campfield, a research leader at Hoffmann–La Roche, the drug company that has bought the rights to Millennium's work on the leptin receptor. In fact, clinical studies by Friedman and others have shown that unlike *obese* and *diabetes* mice, heavy humans generally produce a normal amount of leptin given the amount of fat they are carrying. At least at first glance, there seems to be nothing wrong with their leptin systems.

All of which is no surprise to most obesity researchers, who have long maintained that there must be multiple genes that interact with one another and with economic and psychological pressures to set an individual's susceptibility to weight gain. Although identifying clusters of interrelated genes is considerably trickier than finding single mutations, some labs have made headway in mice. David West of the Pennington Biomedical Research Center in Baton Rouge, La., has been crossing one strain that fattens dramatically on a high-fat diet with a closely related strain that remains relatively lean on the same menu. By tracking the way the trait is passed from one generation to the next, West has proved that the fat sensitivity is carried by one to four dominant genes, and he has narrowed down the chromosome segments on which they could lie. Interestingly, the *tubby* gene happens to rest within one of these segments.

Eventually the genes involved in human weight regulation should be found. But that is the simple part. To make a dent in obesity, physiologists will then have to figure out how all these genes work in real bodies outside the lab. The first step will be to resolve once and for all an old dieters' debate: Do we or do we not have set points—predetermined weights at which our bodies are happiest—and can they be changed?

Set up for Failure

A typical American adult gains about 20 pounds between the ages of 25 and 55. "If you figure that an adult ingests 900,000 to one million calories a year and you calculate the energy cost of those additional 20 pounds," observes Rudolph L. Leibel, co-director of the human metabolism laboratory at Rockefeller, you find that "just a few tenths of 1 percent of the calories ingested are in fact being stored. That degree of control or balance is extraordinary."



SOURCES: New England Journal of Medicine; Annals of Internal Medicine; American Journal of Clinical Nutrition; Journal of the American Medical Association; Circulation

fat from forming would probably work only in children, Evans says, because fat cells in adults usually inflate in size rather than increase in number. But a synthetic molecule that mimics 15d-PGJ_2, called troglitazone, does appear to be an effective drug for the type II diabetes associated with obesity, because it also signals muscle cells to respond normally to insulin.

In mapping the maze of intertwined pathways that control short-term appetite as well as factors (such as fat and carbohydrate levels) that change over days or weeks, researchers are slowly working out how all these signals combine to hold weight steady. Two major theories vie for acceptance: set point and settling point.

The set-point hypothesis is the older and more deterministic. It asserts that the brain continuously adjusts our metabolism and subconsciously manipulates our behavior to maintain a target weight. Although the set point may change with age, it does so according to a fixed genetic program; diet or exercise can move you away from your set point, at least for a time, but the target itself cannot change—or so the theory goes. Last year Leibel and his colleagues Michael Rosenbaum and Jules Hirsch, who are three of the strongest proponents of the set-point theory, completed a study that seems to support their hypothesis.

The physicians admitted 66 people to the Rockefeller hospital. Some of the patients were obese, and some had never been overweight, but all had been at the same weight for at least six months. Over the next three months the subjects ate only precisely measured liquid meals. The doctors ran an extensive battery of tests on the volunteers and then increased the calories that some were fed and put the others on restricted diets. When the subjects had gained 10 percent or lost either 10 or 20 percent of their original weight, the tests were run again to see what had changed.

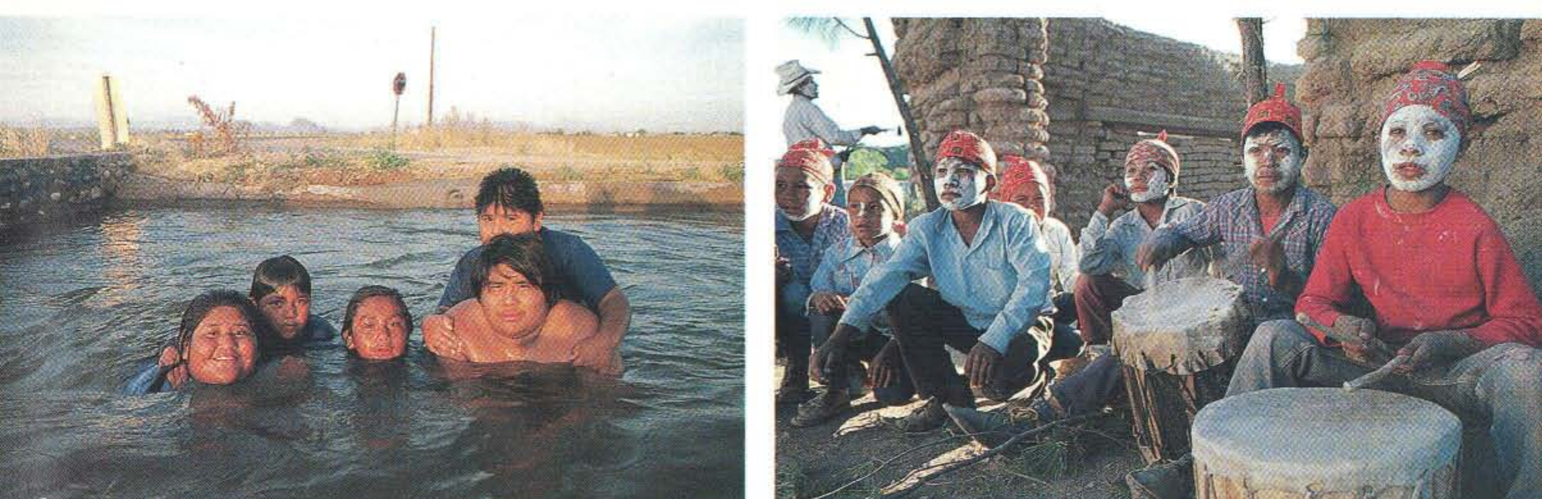
The investigation disproved some tidbits of weight-gain folklore, such as that thin people do not digest as much of their food as heavyweights. The study also found that "the idea that you will be fatter—or will require fewer calories to maintain your starting body weight—as a result of having yo-yoed down and back up again is wrong," Rosenbaum adds. Moreover, the research showed that obese people, when their weight is stable, do not eat significantly more than lean people with the same amount of muscle but less fat.

But the trial's real purpose was to determine how much of a fight the body puts up when people attempt to change the weight they have maintained for a long time—why, in other words, dieters tend to bounce back to where they started. When both lean and obese subjects dropped weight, "it seemed to set off a bunch of metabolic alarms," Leibel recalls. The subjects' bodies quickly start-

Multiple feedback loops maintain the body at a stable weight by shunting messages through the bloodstream and the autonomic nervous system between the brain, the digestive tract, muscle—and, it turns out, fat. Until recently, fat was generally considered just a passive storage tissue. In fact, says Ronald M. Evans of the Salk Institute in La Jolla, Calif., "it is a type of endocrine tissue. Fat secretes signals—hormones such as leptin—and also monitors and responds to signals from other cells."

Last December, Evans reported his discovery of a new hormone, with the catchy name of 15d-PGJ_2, that is produced inside fat cells and seems to trigger the formation of new ones, at least in children. Any drug that tried to interfere with the hormone to prevent new

PHOTOGRAPHS BY JOHN ANNERINO



EFFECT OF ENVIRONMENT on prevalence of obesity is evident in comparisons of Pima Indians in Arizona (*left*) with relatives in Mexico (shown during a holy week festival). Despite a common genetic background, American Pimas have a far higher incidence of obesity—in part because many eat high-fat foods, whereas Mexican Pimas subsist mainly on grains and vegetables.

A Spoonful of Medicine: Obesity Drugs under Development

TISSUES	DRUG	ACTION	DEVELOPER	STATUS
Brain	Dexfenfluramine	Increases the circulation of serotonin, a neurotransmitter that quells appetite	Interneuron with Wyeth-Ayerst Laboratories	Approved by the FDA in April
	Sibutramine	Boosts levels of both serotonin and noradrenaline in the brain, staving off hunger	Knoll Pharmaceutical	Submitted to the FDA for approval in August 1995
	Neuropeptide Y inhibitors	Inactivate NPY, an appetite stimulant that also signals the body to burn more sugars and less fat	Neurogen, Pfizer, Synaptic Pharmaceutical	Phase I trials* began in March
	Bromocriptine	Mimics the neurotransmitter dopamine. Given at certain times of day, may reduce blood sugar and fat production by the liver	Ergo Science	Phase III trials under way for diabetes, planned for obesity
	Leptin	Hormone produced by fat cells and received by receptors in the hypothalamus. Some obese people may be insensitive to leptin; supplemental injections may help	Amgen	Phase I trials began in May
Brain, digestive tract	CCK_A promoters	Increase availability of certain cellular receptors that reduce appetite when stimulated by cholecystokinin (CCK), a family of hormones and neurotransmitters	Astra Arcus USA; Glaxo Wellcome	Preclinical research
	Butabindide	Blocks an enzyme that restores appetite by breaking down CCK. In hungry mice, reduces food intake by 45 percent	INSERM (France)	Preclinical research
Digestive tract	Orlistat	Interferes with pancreatic lipase, one of the enzymes that breaks down fat, so that about one third of the fat eaten passes undigested through the body	Hoffmann–La Roche	Phase III trials complete; FDA application expected by late 1996
	Insulinotropin	Synthetic version of the hormone glucagonlike peptide-1, which may improve obesity-related diabetes by slowing stomach emptying and boosting insulin levels	Novo Nordisk (Denmark)	Phase II trials under way
Fat	Bta-243	Binds to $beta_3$-adrenergic receptor on fat cells, increasing the amount of fat in the blood and burned for energy	Wyeth-Ayerst Laboratories	Preclinical research
Fat, muscle	Troglitazone	Synthetic version of the hormone 15d-PGJ_2, which is produced by fat cells and somehow signals muscle cells to burn fat rather than sugars. May help reverse insulin resistance in obese diabetics	Parke-Davis; Sankyo	Approved in Japan. Phase III trials concluding in U.S.; FDA application expected by late 1996
Entire body	Cytokine regulators	Change the activity of cytokines, hormonelike proteins that act as messengers among cells	Houghten Pharmaceuticals	Phase II trial under way for obesity-related diabetes

*Drugs generally must clear three types of clinical trials before the Food and Drug Administration will approve them for sale. Phase I trials test a drug's safety, and Phase II trials study its effectiveness, both on a small number of patients. Phase III trials must prove that the drug has acceptable side effects and benefits when given to a large group of subjects.

ed burning fewer calories—15 percent fewer, on average, than one would expect given their new weight. Surprisingly, the converse also seems to be true for weight gain. Even rotund people have to eat about 15 percent more than one would expect to stay very far above their set point.

That fact raises a major problem for set-point theory: How does it explain the rapid increase in the prevalence of obesity? "Clearly, set points have to be rising, just as we are getting taller in every generation," Rosenbaum says. "But set points are not changeable in adulthood, as far as we can tell. So there must be a window of opportunity sometime in childhood where the environment influences the set point," he speculates. "If you could figure out when and how that occurs, maybe you could modify the environment then, and you wouldn't have to worry about your kids getting fat 20 years down the line."

That will remain wishful thinking until set-point advocates demonstrate how weight is centrally controlled. Their best guess now, explains Louis A. Tartaglia, a scientist at Millennium, is that "the body's set point is something like a thermostat"—a lipostat, some have called it—and leptin acts like the thermometer.

As you gain weight, Friedman elaborates, "you make more leptin. That shuts off appetite, increases energy expenditure and undoubtedly does other things to restore body weight to the set point. Conversely, if you get too thin, levels of leptin fall, and now you eat more, burn less, and again your weight returns to where it started. Now that we know what the gene and its product are, we can test that simpleminded theory."

Amgen, a biotechnology firm in Thousand Oaks, Calif., that has reportedly promised Rockefeller up to $100 million for the right to produce leptin, has begun injecting the hormone into obese people in clinical trials. "The goal," Rosenbaum says, "is to co-opt your body into working with you rather than against you to maintain an altered body weight" by tricking it into believing it is fatter than it is.

But the body may not be easily fooled. In May, scientists at the University of Washington reported that they had en-

gineered mice that lack the gene for neuropeptide Y (NPY), the most powerful appetite stimulant known. Leptin curtails NPY production; this, it was thought, is how it quells hunger. But mice lacking NPY do not lose weight—something else compensates.

Critics of the set-point hypothesis also protest that it fails to explain the high rates of obesity seen in Nauruans and American Pimas. Moreover, if body fat is centrally controlled, they argue, the amount of fat in your diet should have little impact on your weight. Numerous studies have found the contrary. One recent survey of some 11,600 Scotsmen observed that obesity was up to three times more common among groups that ate the most fat than among those who relied on sugars for most of their energy.

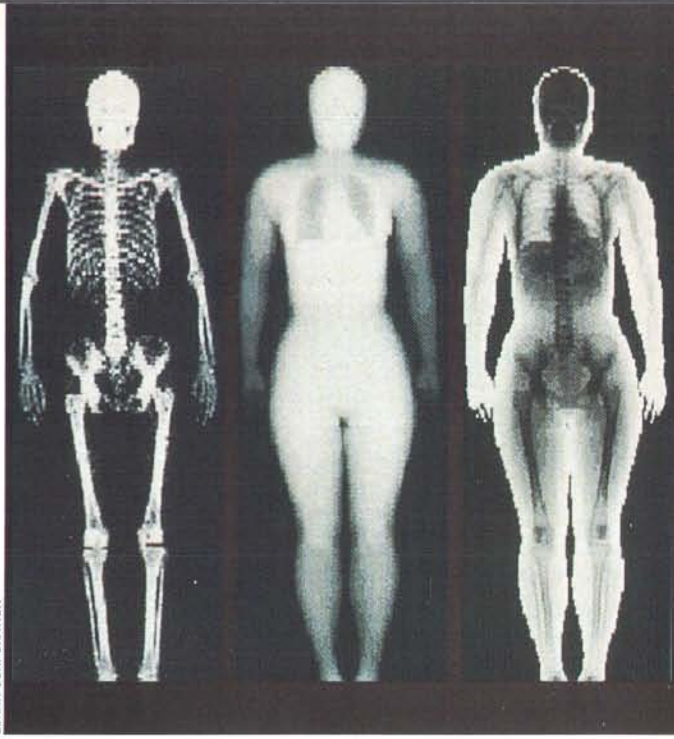
Fat in the Balance

At a conference last year, researchers reviewed the evidence and judged that although the set-point hypothesis has not been disproved, there is more "biological merit" to the idea of a "settling point." This newer theory posits that we maintain weight when our various metabolic feedback loops, tuned by whatever susceptibility genes we carry, settle into a happy equilibrium with our environment. Economic and cultural changes are upsetting this equilibrium and propelling more people—those with more genetic risk factors—into obesity.

The prime culprit suspected in this trend is hardly surprising: it is the fat dripping off hamburgers, smoothing out ice cream and frying every meat imaginable. But biochemists are at last working out precisely why fat is bad. For years, they have known that people fed a high-fat meal will consume about the same amount as those given a high-carbohydrate meal. Because fat has more calories per bite, however, the subjects with greasy grins tend to ingest more energy than they can burn, a phenomenon known as passive overconsumption.

One reason for this, according to biopsychologist John E. Blundell of the University of Leeds, seems to be that the systems controlling hunger and satiety respond quickly to protein and carbohydrates but slowly to fat—too slowly to stop a high-fat meal before the body has had too much. Metabolic systems seem to favor carbohydrates (which include sugars and starches) as well. Knock down a soda or a plate of pasta, and your body will soon speed up its carbohydrate combustion. Polish off a bag of pork rinds, however, and your fat oxidation rate hardly budges, points out Jean-Pierre Flatt, a biochemist at the University of Massachusetts Medical School. Most incoming fat is shipped directly to storage, then burned later only if carbohydrate reserves dip below some threshold, which varies from person to person.

There is another way to increase the rate at which fat is burned for energy: pack on the pounds. More fat on the body yields more fatty acids circulating in the bloodstream. That in turn boosts fat oxidation, so that eventually a "fat balance" is reached where all the fat that is eaten is combusted, and weight stabilizes. Many genetic and biological factors can influence the fat oxidation rate and thus affect your settling point in a particular environment.



LUNAR CORPORATION

CLEARER PICTURE of body fat (*right*) is obtained by mathematically combining images taken with high-energy (*left*) and low-energy (*center*) x-rays.

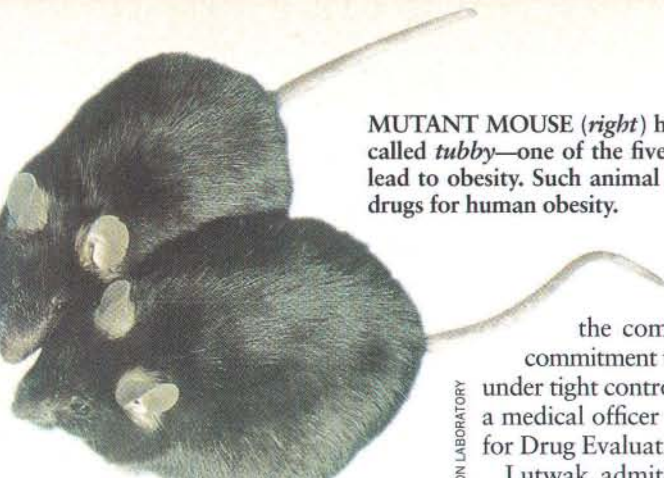
Olestra, an artificial fat approved earlier this year by the Food and Drug Administration, may change that rate as well. Olestra tastes more or less like an ordinary fat, but it flows undigested through the body. A preliminary study by George A. Bray, Pennington's executive director, suggests that the ingredient may short-circuit passive overconsumption. For two weeks, Bray replaced the natural fat in his subjects' meals with olestra. "They did not compensate at all by eating more food," he reports, adding that "it remains to be seen whether that holds up in longer-term studies."

The fat balance explains in part why settling points vary among people who overeat fat: some oxidize fat efficiently at normal weights; others burn too little until excess pounds force the oxidation rate up. But the model does not by itself explain why some do not overeat at all. To answer that, Flatt has proposed a "glycogen hypothesis."

The human body can store about a day's supply of carbohydrates in the form of glycogen, a simple starch. Glycogen reserves function somewhat like fuel tanks; we partially refill the stores with each meal but rarely top them off. In fact, the range between "empty" and "full" appears to be a matter of individual preference, influenced by such factors as the diversity and palatability of food at hand, social pressures and meal habits. People who are content with lower glycogen levels or who frequently deplete them through exercise burn fat more readily than those who like to keep their tanks full, Flatt suggests. But he concedes that the "crucial link from glycogen stores to appetite remains to be proven."

Researchers need more evidence before they can pronounce either set point or settling point—or neither—correct. James O. Hill of the University of Colorado Health Sciences Center has begun collecting some of those critical data. He is assembling a registry of the most precious resource in obesity research: the people who have lost a large amount of weight and kept it off for several years without relapse. Hill has already identified about 1,000 such individuals and has begun examining a handful for biochemical clues to their success.

Unfortunately, no current explanation of weight regulation leaves much room for voluntary control; all the metabolic cycles involved are governed subconsciously. Settling-point theory does at least suggest that sufficiently drastic changes in lifestyle might prod the body to resettle at a new weight. But without assistance, changes radical enough to make a difference are evidently uncom-



MUTANT MOUSE (*right*) harbors a flaw in a gene called *tubby*—one of the five discovered so far that lead to obesity. Such animal models may yield new drugs for human obesity.

JACKSON LABORATORY

fortable enough to be infeasible—for millions of dieters have tried this strategy and failed.

Getting over the Hump

Increasingly, obesity researchers argue that the most effective assistance they can provide their patients will probably be pharmacological. "The treatment philosophy of the past 40 years, which has been to train patients to eat differently, is simply not going to cure the epidemic of obesity that we see worldwide," asserts Barbara C. Hansen, director of the obesity research center at the University of Maryland School of Medicine.

Untangling the biology beneath body fat has created a plethora of new drug targets that has drawn dozens of pharmaceutical firms off the sidelines [*see table on page* 74]. The potential market is enormous, not only because obesity is common and growing but also because even an ideal drug will have to be taken indefinitely, according to Hansen and others. "Obesity isn't curable," Bray says. "It's like high blood pressure. If you don't take the medication, your blood pressure won't stay down. And if you don't take drugs—or do something—to treat obesity, your weight won't stay down."

Part of the reason for the resurgence of commercial interest is a shift in policy at the FDA, which decided in May to allow the appetite suppressant dexfenfluramine to be prescribed for obesity in the U.S., as it already is in 65 other countries. It is the first weight-loss drug approved in the U.S. in 23 years, and nearly all obesity researchers agree it has been too long coming. The FDA also recently relaxed its guidelines for obesity-drug applications. "As our compromise right now, we're suggesting that a company can present us with two years of data—in some cases, one year if the data look good enough and the company gives us a firm commitment to do follow-up studies under tight controls," says Leo Lutwak, a medical officer with the FDA's Center for Drug Evaluation & Research.

Lutwak admits that with only two years of information, the FDA may approve drugs that turn out to have serious long-term side effects. "The best we can hope for is something like insulin for the treatment of diabetes," Leibel says. Insulin rescues a type I diabetic by replacing a hormone that is missing. "But after 15 years, you begin to have complications of our inability to perfectly mimic the biology," Leibel continues. "If we're lucky, that's the kind of problem we'll face in the treatment of obesity." Lutwak responds that "when that happens, the public will be informed, and they will have to make a decision about whether it is worth it."

If the long-term cost of treatment is unknown, the benefits are becoming clearer, thanks to studies on people who have an operation, called gastroplasty, that reduces the size of the stomach. Although infrequently used in the U.S., the procedure has proved remarkably effective in Sweden. A long-term study there of 1,150 obese patients who underwent gastric surgery found that they typically dropped 66 pounds over two years—88 pounds if a more severe procedure was used—whereas control subjects given standard dietary treatment lost nothing. The surgery cured more than two thirds of those with diabetes, compared with 16 percent cured in the control group. Likewise, twice as many (43 percent) of the hypertension cases were cured by the operation.

Gastroplasty has drawbacks in addition to the risks that always accompany major surgery—principally a high rate of digestive complications. Drug treatments might be better, but Hansen's work with rhesus monkeys suggests that prevention would be best. A decade ago her team began a trial on young adult monkeys, equivalent in maturity to 20-year-old men. The researchers adjusted the animals' food supply so that they neither gained nor lost weight. "In the past 10 years we have had 100 percent success preventing both obesity and type II diabetes," Hansen asserts. "In the control group, which was simply allowed to feed freely on the same diet, half are diabetic. Because everything we know about human obesity is also true of nonhuman primate obesity, that shows you the power of weight control."

It does not, unfortunately, demonstrate a feasible way to achieve it. The NIDDK has launched a program to educate Americans about ways to avoid weight gain, but Susan Z. Yanovski, the program's director, admits that so far it has had little perceptible impact. There is no major lobbying organization for the disease, notes Pi-Sunyer, and the NIH directs less than 1 percent of its research funding at obesity. "Many people seem to be unaware of how big a health problem this is now and how big it is going to grow, particularly when you look at the increasing obesity of children," Yanovski says. Because obese adolescents usually become fat adults, "we're really heading for trouble in another 20 to 30 years," she adds.

At least one grade school intervention has had modest success, knocking a few percentage points off the number of children who turn into overweight adolescents by taking fat out of the children's lunches, giving them more strenuous recreation and educating their parents about weight control. "We have to be very careful about putting children on restrictive diets," Yanovski warns. "That is inappropriate. But we can be more proactive in getting our kids away from the television set, more physically active, riding their bikes instead of being driven everywhere. If people recognize that this is a serious public health problem affecting their children, then maybe they will start taking some action." If not, economists should start adjusting their models now to account for the tremendous health care cost increases that lie ahead. SA

Further Reading

WEIGHING THE OPTIONS: CRITERIA FOR EVALUATING WEIGHT-MANAGEMENT PROGRAMS. Edited by Paul R. Thomas. National Academy Press, 1995.

REGULATION OF BODY WEIGHT: BIOLOGICAL AND BEHAVIORAL MECHANISMS. Edited by C. Bouchard and G. A. Bray. John Wiley & Sons, 1996.

Additional information, including an extensive bibliography, is available on the *Scientific American* World Wide Web site at http://www.sciam.com

NEW ADVENTURES & OLD FAVORITES **SCIENTIFIC AMERICAN** FOR THE DISCRIMINATING SHOPPER!

SELECTIONS

DISCOVER A NEW WAY OF LOOKING AT THE ELEMENTS with our unique ***Periodic Round Table***. Observe the elements that make up the universe in a new dimension. This symmetrical and elegant piece allows a much more direct appreciation of the electronic structure of atoms than any traditional two-dimensional table of elements could.

This brilliant design is perfect for encouraging students to develop creative new ideas about the elements and scientific models in general.

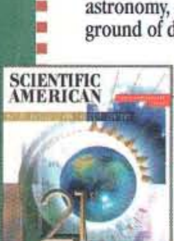
And your colleagues will be amazed not only by its novelty when they see it on your desk but also by its quality both as a tool of science and its durable construction. Each of the eight pieces is naturally finished, solid hardwood with permanently branded lettering that will stand up to the heavy usage in a classroom or lab.

Periodic Round Table $49.95

MAKE A FASHION STATEMENT. SCIENTIFIC AMERICAN commissioned this 100% silk, hand-finished, designer-quality ***Sciences Necktie*** with striking images representing physics, astronomy, anthropology, chemistry and biology against a background of deep blue. Gift-boxed.

Sciences Necktie $34.95

Just $30.00 for each additional tie.



SCIENTIFIC AMERICAN SHEDS NEW LIGHT ON ***Key Technologies for the the 21st Century.*** This latest single-topic subscription issue reveals the future of intelligent software, virtual reality, space travel, artificial organs, sustainable agriculture and superconductors, gene therapy and other promising technologies for the next century. Take a remarkable peek into our technological future.

Key Technologies for the 21st Century $4.95

STEPHEN HAWKING'S interactive Mac and Windows CD-ROM, ***A Brief History of Time*** takes readers through space and time to distant galaxies, black holes, and alternate dimensions to confront the supreme questions about the origin and nature of the universe. From the great Renaissance thinkers to modern astrophysics, Hawking guides you with wit and charm on your incredible journey using film excerpts, spectacular new images and interactive animations. Bonus: the book's complete text is included.

A Brief History of Time CD-ROM $49.95

This CD-ROM is compatible with Macintosh and Windows computers Minimum system requirements: Macintosh—System 7; 6MB RAM; 13" monitor displaying 256 colors; CD-ROM drive with sustained transfer rate of 150Kbps. ibm compatibles—486/33MHz CPU; 8MB RAM; DOS 5.0; Windows 3.1; VGA+ monitor displaying 256 colors at 640x480 resolution; CD-ROM drive with sustained transfer rate of 150Kbps (double-speed recommended); sound card.

B

C

D

PUZZLE LOVERS BEWARE! Our latest collection of mesmerizing full-color jigsaws is the ultimate challenge. Each is 19.6" x 27.5" full-color, 1000 pieces.

CREATED TO CELEBRATE OUR 150TH ANNIVERSARY, we dare you to try the stunning new ***Anniversary Fractal Puzzle.*** Fractals can be as maddening as they are mesmerizing. This one is based on the formula $Z=Z^2+c$.

THE ELUSIVE BEAUTY of ***Spektrum's Fractal Puzzle*** is revealed by Benoit Mandelbrot's set for the function, $f(z)=z \tan z+c$, which includes the popular apple-like pattern he made famous. But forget the math. Get this dazzling puzzle just for the fun of it.

TAKE A STEP BACK to view the Earth from aboard the space shuttle, Endeavor. This radar image reveals an impressive view of the erupting ***Kliuchevskoi Volcano*** in Kamchatka, Russia. Piece it together and view the home planet in a new light.

A

ZOOM IN CLOSE TO DISCOVER the beautiful hidden world of photomicrography. Magnified 250 times, this close-up of ***Crystals under Polarized Light*** is an extraordinary image to challenge any puzzle enthusiast.

Anniversary Fractal Puzzle $29.95 (A, above)
Spektrum's Fractal Puzzle $29.95 (B, above)
Kliuchevskoi Volcano Puzzle $29.95 (C, above)
Crystals under Polarized Light Puzzle $29.95 (D, above)

Just $29.00 for each additional puzzle.

THE ULTIMATE T-SHIRT is now available, exclusively from SCIENTIFIC AMERICAN. You'll find it easy to take your fascination with science wherever you go when you wear this bright and beautiful ***Fractal T-Shirt*** of pre-shrunk 100% cotton. Like the puzzle, this T-shirt sports a portion of the fractal pattern, computer-generated from the equation $Z=Z^2+c$. This bold, full-color design is sure to turn heads everywhere.

Fractal T-Shirt large $18.95
Fractal T-Shirt x-large $18.95

▲ *Photocopy order form or cut along dotted line and mail in an envelope with payment.* ▲

HERE'S MY MANIFEST. START MY ADVENTURE.

Name *(print clearly)*

Address Apt.

City/Province

State/Country Zip/Postal Code

Method of payment (no cash or c.o.d.'s, please):

☐ Check ☐ Money order
☐ Visa ☐ MasterCard ☐ Access ☐ Eurocard ☐ AmEx

Card number Expiration date

Signature

☐ Please send me a free catalog. RP896

Item	*Quantity*	*Price each*	*Price each 2 or more*	*Item sub-total*	*S&H, each*	*Total*
Periodic Round Table		$49.95	—		$5.00	
Sciences Necktie		$34.95	$30.00		$3.95	
Key Technologies		$4.95	—		Free	
Brief History of Time		$49.95	—		$5.00	
Spektrum Fractal Puzzle		$29.95	$29.00		$3.95	
Anniversary " "		$29.95	$29.00		$3.95	
Volcano Puzzle		$29.95	$29.00		$3.95	
Crystals Puzzle		$29.95	$29.00		$3.95	
Fractals T-Shirt Large / X-large		$18.95	—		$3.95	
					Total order:	
					Sales tax:*	
					Grand total:	

Double shipping charges outside the US and Canada. Call 515-246-6911. Or fax 515-246-6933. Or mail your order to:
Scientific American Selections
PO Box 11314; Des Moines, IA 50340-1314

*Residents of CA, IA, IL, MA, MI, NY please add appropriate sales tax.
Canadian residents please include GST. GST no. R127387652.
Quebec residents please include QST. QST no. Q1015332537.
Payment accepted only in US funds drawn on a US bank.

PLACE YOUR ORDER TODAY!

OR CALL 1-800-777-0444 AND START YOUR ADVENTURE NOW!

by Shawn Carlson

Detecting Micron-Size Movements

There's a lot going on down among the microns. What we perceive as a rigid surface squashes easily under a finger's gentle pressure when viewed from a distance of a millionth of a meter. Increasing the temperature sends objects at that scale into even more violent upheavals.

Biological processes reshape many living things on this scale. For example, every beat of an insect's dorsal vessel—essentially, its heart—flexes its abdomen by a few microns.

Now, thanks to John R. B. Lighton, a biologist at the University of Nevada, these tiny movements can be readily detected. (Lighton is not only a world-renowned physiologist but also a kindred spirit to amateur scientists everywhere, always striving to find the most direct and least expensive solution to vexing experimental challenges.) He realized that by detecting the microscopic flexings of an insect's body, he could in effect put a tiny stethoscope on the creature. This technique opens all kinds of micromotions for study, including the slight distortion of materials caused by changes in ambient temperature and pressure.

Lighton's ingenious method allows experimenters to embark on a fantastic voyage into the microscopic universe. Now anyone can detect movements as small as half a micron—about the wavelength of visible light—for less than $40.

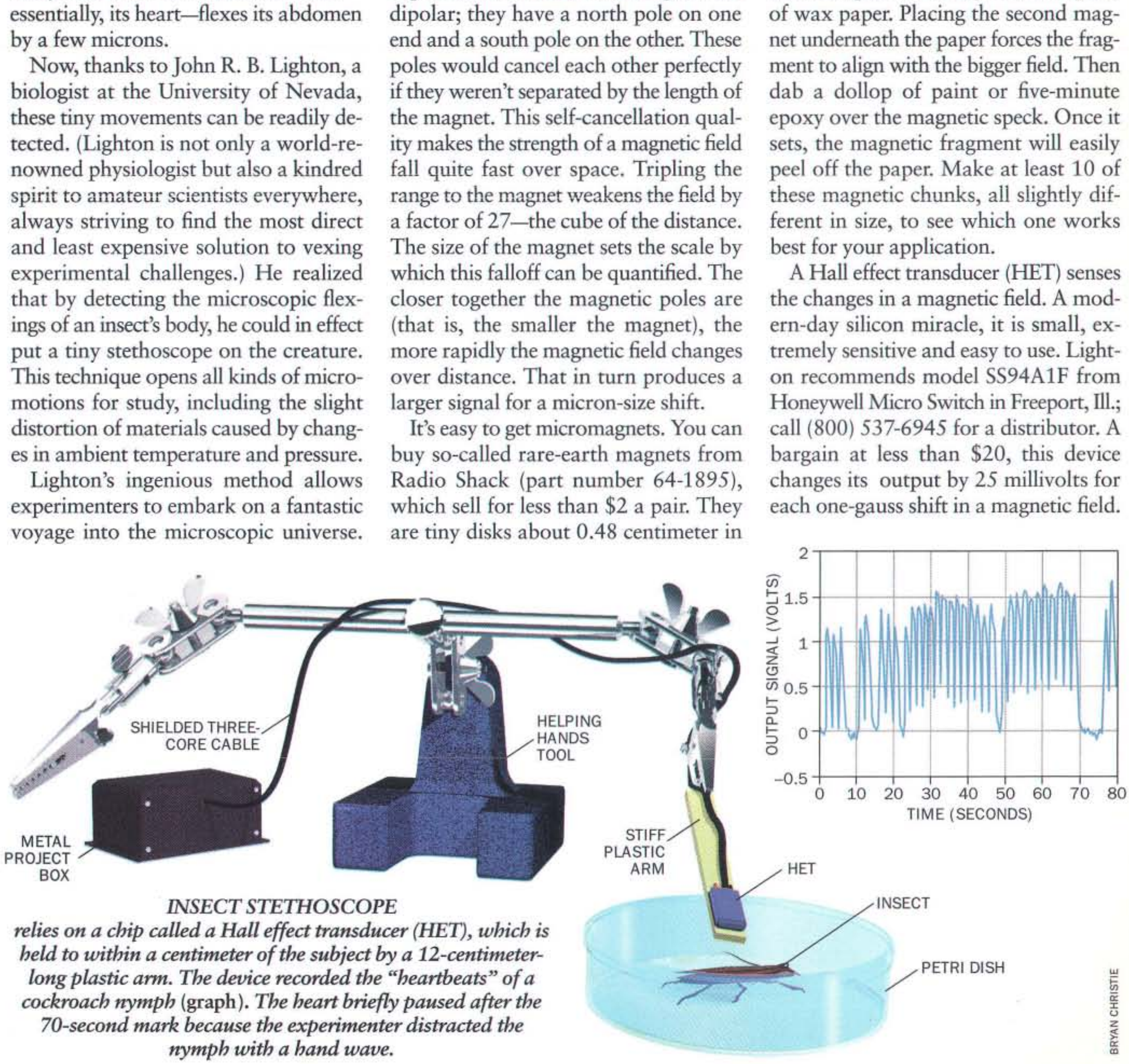
Lighton senses micromotions by using minuscule magnets that he attaches to the moving objects. He then relies on a special sensor that picks up the variations in the magnetic field caused by the shifting magnet.

The sensitivity of Lighton's detector depends on the fact that all magnets are dipolar; they have a north pole on one end and a south pole on the other. These poles would cancel each other perfectly if they weren't separated by the length of the magnet. This self-cancellation quality makes the strength of a magnetic field fall quite fast over space. Tripling the range to the magnet weakens the field by a factor of 27—the cube of the distance. The size of the magnet sets the scale by which this falloff can be quantified. The closer together the magnetic poles are (that is, the smaller the magnet), the more rapidly the magnetic field changes over distance. That in turn produces a larger signal for a micron-size shift.

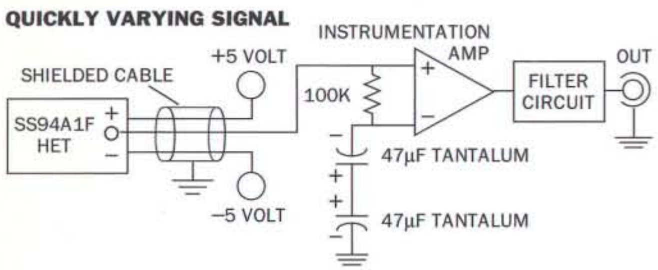
It's easy to get micromagnets. You can buy so-called rare-earth magnets from Radio Shack (part number 64-1895), which sell for less than $2 a pair. They are tiny disks about 0.48 centimeter in diameter and 0.16 centimeter high ($^3/_{16}$ by $^1/_{16}$ inch). At the surface, the magnetic field, which is oriented perpendicular to the flat part of the disk, is about 20,000 times that of the earth.

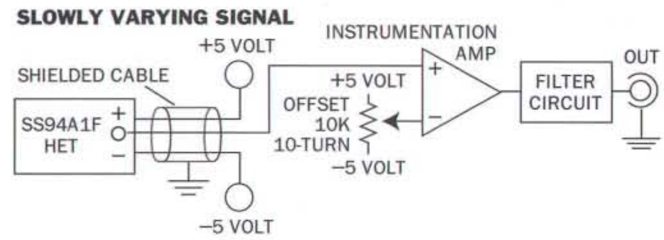
If these magnets are too big for your project, then crush one with a pair of pliers. Made from a brittle ceramic, they will shatter into little shards. You need to make sure, though, that you know the direction in which the magnetic field of these shards points. Using nonmagnetic tweezers, position a fragment on a piece of wax paper. Placing the second magnet underneath the paper forces the fragment to align with the bigger field. Then dab a dollop of paint or five-minute epoxy over the magnetic speck. Once it sets, the magnetic fragment will easily peel off the paper. Make at least 10 of these magnetic chunks, all slightly different in size, to see which one works best for your application.

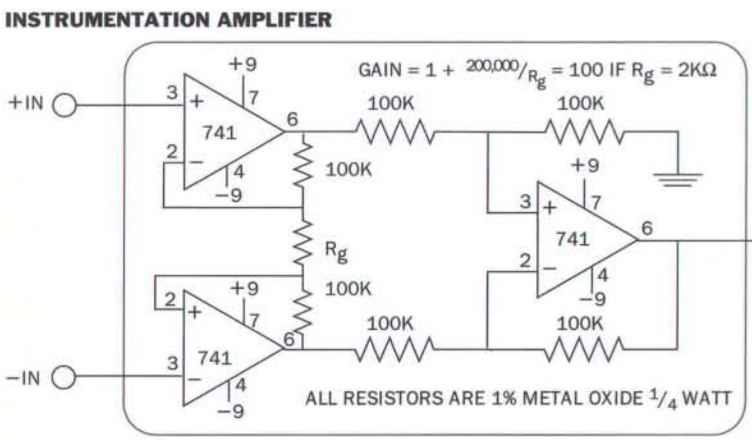
A Hall effect transducer (HET) senses the changes in a magnetic field. A modern-day silicon miracle, it is small, extremely sensitive and easy to use. Lighton recommends model SS94A1F from Honeywell Micro Switch in Freeport, Ill.; call (800) 537-6945 for a distributor. A bargain at less than $20, this device changes its output by 25 millivolts for each one-gauss shift in a magnetic field.

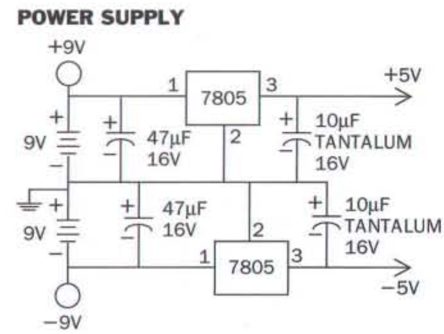


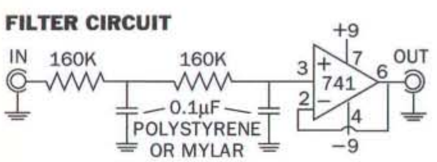
INSECT STETHOSCOPE
relies on a chip called a Hall effect transducer (HET), which is held to within a centimeter of the subject by a 12-centimeter-long plastic arm. The device recorded the "heartbeats" of a cockroach nymph (graph). *The heart briefly paused after the 70-second mark because the experimenter distracted the nymph with a hand wave.*











JOHNNY JOHNSON

RATE OF MOTION dictates the necessary circuitry. If the signal changes much over about 30 seconds, choose the quickly varying circuit. For more leisurely signals, use the slowly varying circuit. The instrumentation amplifier can be constructed from three operational amplifiers. A filter circuit and a power supply complete the system.

Secure your HET no more than one centimeter away from your subject. For instance, if you are monitoring insects, you can epoxy the HET to a rigid piece of plastic and hold it above the subject with a device called Helping Hands, a soldering aid sold by Radio Shack.

A HET records all magnetic fields, including the earth's. This indiscriminateness means that the detector will always produce a large constant voltage signal (created by the earth and the magnet). On top of this voltage constant will be the small changing signal created by the magnet's motion.

You can forget about trying to boost the signal with a single operational amplifier (op-amp). A single op-amp cannot accurately amplify a small signal on top of a voltage constant. What you need is an instrumentation amplifier. Like op-amps, instrumentation amplifiers are available as inexpensive, integrated circuits. Entry-level devices cost about $5; the Cadillacs of these chips sell for about $20. The AD524 from Analog Devices in Norwood, Mass., works well; to order, call (800) 262-5643, extension 3. You can also construct an instrumentation amplifier from three type 741 op-amps [*see illustration on preceding page*].

If you're monitoring temperature or another signal that varies slowly, use Lighton's slowly varying signal rendition of the circuit. For flexing insect abdomens and other activities that change significantly over 30 seconds or so, use the quickly varying signal circuit. The circuit employs a clever technique that should be in every amateur's (and professional's) tool kit.

The trick begins by splitting in two the voltage from the HET. One signal goes into the amplifier's positive input. The other goes into a low-pass filter that only passes signals that oscillate slower than about one cycle every 30 seconds. Because an insect's heart contracts in much less time, the filter strips out that signal and passes the large constant voltage (the DC offset). This filtered voltage is then fed into the instrumentation amplifier's negative input. An instrumentation amplifier boosts the difference between its two inputs, so the troublesome offset voltage is automatically subtracted, leaving only the coveted signal.

Signal wires can introduce extraneous signals. They act like antennae, picking up electromagnetic energy, such as emanations from 60-cycle power lines, and then dumping it directly into your amplifier. To minimize the effect, keep the leads between the HET and the amplifier short. Additionally, you should use shielded wire. Lighton relies on three-core shielded cables. An electronics store may stock them, or you can make your own. Twist together three different colored wires, one each for the positive, negative and signal leads of the HET. Wrap the wires inside a layer of aluminum foil and connect the foil to the circuit's ground with a short wire. For protection, add a layer of duct or electrical tape around the foil. The filter circuit provides another barrier to power-line noise. Finally, encase all your electronics inside a grounded metal project box.

You can read the output with a digital voltmeter or, better yet, use an analog-to-digital converter to record the data into a computer. Several software

Continued on page 83

MATHEMATICAL RECREATIONS

by Ian Stewart

Shedding a Little Darkness

Angela is standing in a room with perfectly reflecting walls. Somewhere in this Hall of Mirrors, her friend Bruno lights a match. No matter what the shape of the room or where the two are located, can Angela always look around and see Bruno's match or its reflection? Or equivalently, does the light from the match fill the whole room, not missing a single point—no matter where the match is placed?

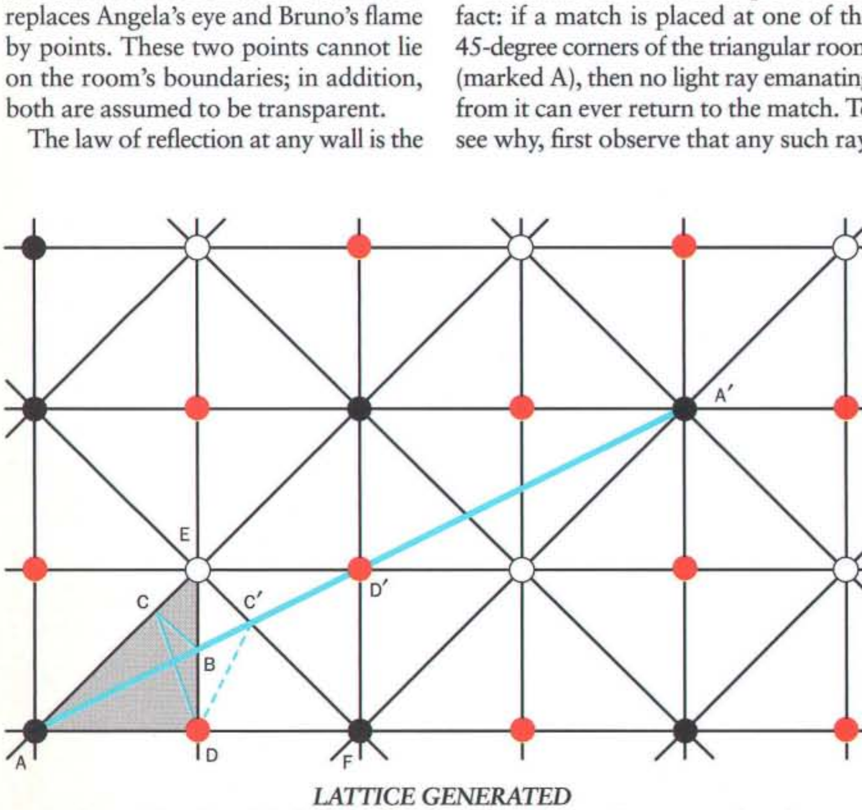
This problem was first asked in print by Victor Klee in 1969, but its origins are thought to go back farther, at least to Ernst Straus in the 1950s. It comes in several variants. The room may be two- or three-dimensional (if the latter, then its floor and ceiling must also be mirrors). It may have flat walls—being polygons in two dimensions, polyhedrons in three—or curved walls. In all versions, the standard mathematical idealization replaces Angela's eye and Bruno's flame by points. These two points cannot lie on the room's boundaries; in addition, both are assumed to be transparent.

The law of reflection at any wall is the usual one: "angle of incidence equals angle of reflection" (see also the July column on shoelaces). Wherever these angles are ill defined—such as at an edge or a vertex—one assumes that the light ray is absorbed and travels no farther.

The answer to the question for plane polygonal rooms was published by George W. Tokarsky in the December 1995 issue of the *American Mathematical Monthly* (Vol. 102, No. 10). Tokarsky's elegant proof appropriately involves a "reflection trick" and, like all the best mathematics, is astonishingly simple.

The key idea is to start with an isosceles right triangle—a square divided in half along a diagonal. Such a triangle, marked AED [*see illustration below*] can be "unfolded" into a regular lattice pattern by repeatedly reflecting it about its three sides.

The lattice is first used to prove a key fact: if a match is placed at one of the 45-degree corners of the triangular room (marked A), then no light ray emanating from it can ever return to the match. To see why, first observe that any such ray,



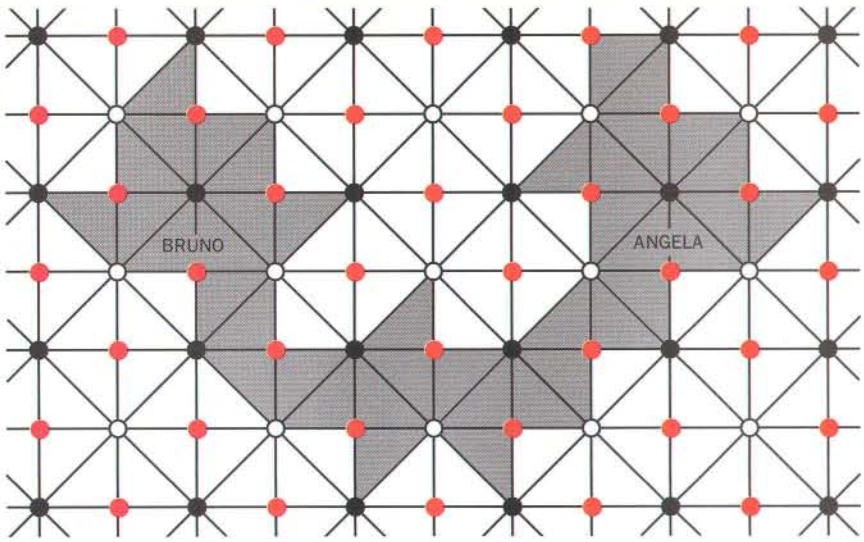
LATTICE GENERATED ***by a right triangle holds the key to understanding reflections within the mirror room AED. A light ray emanating from A and bouncing around within the room can be "unfolded" into a straight line in the lattice. Where this line ends reveals the fate of the original ray.***

such as ABCD, can be unfolded in the same manner as the triangle itself. For instance, the segment BCD inside the triangle is reflected about the wall ED to become BC′D on the other side of the wall. Further, on reflection about EF, C′D unfolds to C′D′. So ABCD unfolds to give ABC′D′. Note that ABCD terminates at D because that is a corner of the triangle; equivalently, the "unfolded" point D′ lies at a point of the lattice. The law of reflection implies that ABC′D′ is a straight line, a fact that is crucial to what follows.

I have colored the three vertices of the triangle so that the vertex A at a 45-degree angle is black, the vertex E is white and the 90-degree vertex D is red. Each lattice point is also colored, according to the vertex that falls on it as the triangle is unfolded. I will argue that if there was a path leading from A back to A, this path must first hit a red or white corner—where the ray is absorbed.

To prove this statement, imagine that we have some path from A to A. Unfold this path, following the pattern of unfolding triangles; the process ensures that the unfolded ray is a straight line terminating at a black lattice point A′. Now, these black dots are spaced an even number of lattice units apart in both the horizontal and the vertical directions: their coordinates are even integers. Thus, the midpoint of AA′ has an odd integer for at least one coordinate—and is therefore a red or white lattice point.

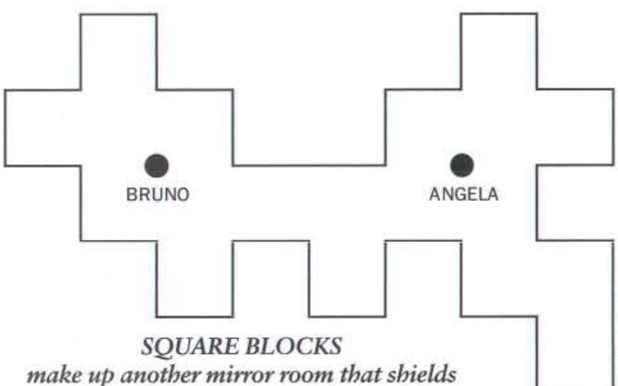
This argument fails, however, if both coordinates are multiples of four. But then the midpoint A″ of AA′ has even coordinates and is a black lattice point, so we can try again for A′A″. Either the midpoint of that is a red or white lattice point, or A″ also has coordinates that are both multiples of four. If the latter, we can replace A″ by the new midpoint A‴, and so on. After a finite number of such replacements, we must ultimately come to a segment that has an odd-integer coordinate. For example, if the coordinates of A′ are 48 horizontally and 28 vertically, A″ has coordinates (24,14) and A‴ has coordinates (12,7). The midpoint at A‴ is therefore a red or white lattice point.



MIRROR ROOM
made from a right-triangular lattice ensures that no light from Bruno reaches Angela.

Thus, any unfolded path that joins A to a black lattice point must hit a gray or white lattice point. So if we fold the path up again, we conclude that the original path hits one of the other two corners before it gets back to A.

We can construct polygonal rooms by fitting together horizontal, vertical or diagonal segments of this triangular lattice [*see illustration above*]. Suppose that a light ray bounces around inside such a room, starting at Bruno and ending at Angela—both of whom are at black points. Then we can fold that ray up to get a path within a single triangle that generates the lattice. But one can establish that any path between two black points will hit a red or white vertex, so unfolding again we conclude that the original ray must also hit a red or white dot. Suppose we arrange for the following three conditions to hold:



SQUARE BLOCKS
make up another mirror room that shields Bruno's light from Angela.

- The two black dots representing Angela and Bruno are in the interior of the room.
- No red or white dot lies in the interior.
- Every red or white dot on the boundary of the room lies at a vertex.

Then any ray that hits a red or white dot must hit a vertex and be absorbed—so there is no such light ray at all.

If you try to design such rooms, you will find that it takes a certain amount of ingenuity. For example, you have to add extra triangles to introduce additional bends into the boundary; unless you are careful, these twists may create extra interior lattice points that violate the second condition.

The room illustrated is built from 39 reflected copies of an isosceles right triangle; Tokarsky's article includes one with 29 component triangles. Can you find this room, or others? Tokarsky also develops a similar theory for rooms obtained by unfolding a square or by using triangles of other shapes, as well as three-dimensional rooms obtained from similar principles.

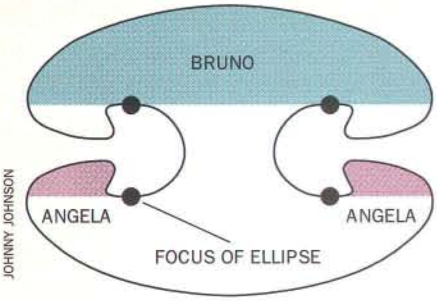
These examples show that in a polygonal room there may be positions where a match does not illuminate every point in the room. All we have proved, however, is that at least one point is unlit. Is it possible for there to be a whole region of nonzero area that is not illuminated? This problem is distinctly trickier. For all we know, rays starting from Bruno can pass as close as we wish to Ange-

FEEDBACK

The April column "How Fair Is Monopoly?" generated the largest mailbag ever. Most readers pointed out that the analysis ignored many features of the real game, notably the "Go to Jail" square but also the Community Chest and Chance cards. Some readers assumed that I didn't know the rules of the game, whereas others deduced (correctly) that I had made some simplifying assumptions but did not make them explicit. One reader wondered if the article was an April Fools' spoof: no, not an intentional one!

A few readers provided the missing analysis of the realistic game. I wish to acknowledge in particular Stephen Abbott of Northfield, Minn.; William J. Butler, Jr., of Portsmouth, R.I.; Thomas H. Friddell of Maple Valley, Wash.; Earl A. Paddon of Maryland Heights, Mo.; and David Weiblen of Reston, Va. Together, they have taught me a great deal—so much so that I intend to report their results, and any others that come to my attention, in a full column.

In reality, the "Go to Jail" square definitely produces a skewed probability distribution. The jail square itself is the most frequently visited, with a probability of 5.89 percent as compared with the "equidistributed" value of 2.5 percent (or 2.44 percent if "Just Visiting" and "Jail" are distinguished, which seems sensible). The next most likely square to be landed on is Illinois Avenue, with a probability of 3.18 percent. The square least often visited is the third "Chance" square around from "Go," probability of 0.871 percent—apart from "Go to Jail," which is not actually visited at all, because you're off to the pokey. *—I.S.*



ELLIPTICAL CURVES *outline a room with regions where Angela is shielded from Bruno's light.*

la—we have only proved that they cannot hit her head on.

The answer for polygonal rooms seems unknown. But Roger Penrose, with a collaborator, showed in 1958 that if a room has curved sides, unlit regions can exist. For example, recall that an ellipse has two special points, or foci. It can be proved that any light ray that passes between the two foci and bounces off the curve will again cross the straight line joining them before it next hits the curve. Bearing this property in mind, one can easily check that a room made from two halves of an ellipse [*see illustration above*] has unilluminable regions. Specifically, rays originating in the blue shaded region (Bruno) can never enter the pink one (Angela).

There are many similar problems, some solved, some not. You can find a selection in *Unsolved Problems in Geometry*, by Hallard T. Croft, Kenneth J. Falconer and Richard K. Guy (Springer-Verlag, 1991), and in *Old and New Unsolved Problems in Plane Geometry and Number Theory*, by Victor Klee and Stan Wagon (Mathematical Association of America, 1991). For instance, Jeffrey B. Rauch of the University of Michigan has shown that there is a curved room, with a smooth boundary at every point except one, that requires infinitely many matches to illuminate fully.

Rauch has also proved that for any finite number of matches there is a smoothly curved room that cannot be illuminated by that many matches. And Janos Pach of the City College of New York has asked this very elegant question: If you light a match in a forest of perfectly reflecting trees, must the light be visible from outside? The trees can, for example, be modeled as circles, with the problem being posed in the plane. No one knows the answer. SA

The Amateur Scientist,
continued from page 80

packages that link the signals to your computer are available [see "The New Backyard Seismology," Amateur Scientist, April]. Use shielded coaxial cable for the output connections, to prevent the HET from detecting the signal.

Lighton obtained some remarkable results after he superglued a whole rare-earth magnet to the abdomen of a *Blaberus discoidalis* nymph, a relative of the American cockroach. With the instrumentation amplifier's gain set to 100, the signal caused by the contractions of the dorsal vessel—the insect's heartbeats—is striking. After about 70 seconds of recording data, Lighton waved his hand in front of the nymph. The insect's heart stopped beating for several seconds. According to Lighton, that happened because the creature's nervous system may be too limited both to maintain circulation and to attend to stimuli.

Of course, any crawling by the insect will disrupt your results, so record data only when it is still. If the insect moves, it will generate a huge voltage signal that jumps well off the scale. In fact, Lighton reports that large signals occur whenever the insect opens its spiracles to breathe, about once every five to 30 minutes. By lowering the gain of your instrumentation amplifier, you can also monitor insect respiration.

Recording the vital functions of insects is just one of the experiments you can do. By attaching the magnet to the bottom of a heavy pendulum and affixing the HET to the floor one centimeter below the magnet, you can make an extremely sensitive seismograph. It can significantly extend the lower range of an amateur seismography station, as described in the April column.

By connecting the magnet to a sheet of Mylar stretched tightly over the mouth of a jar, you might record the atmospheric pressure that pushes against the membrane. Other suggestions appear on the World Wide Web site of the Society for Amateur Scientists. I invite you to invent, experiment and discover—and let me know what you find.

For more information about this project, send $5 to the Society for Amateur Scientists, 4951 D Clairemont Square, Suite 179, San Diego, CA 92117, or download it at http://www.thesphere.com/SAS/ SA

"The Girl with a Pearl Earring," from Mauritshuis, the Hague

THE CULTURE MACHINE: SCIENCE AND ART ON THE WEB

Review by Ben Davis

What distinguishes formerly classified digital satellite image archives from abstract art? As I gaze at my computer screen I can't always tell the difference. The line between art and scientific artifact has become a blur as cultural and scientific institutions move onto the Internet, suddenly sharing the same on-line formats and virtual architectures. In the process, the World Wide Web seems to be creating a broad definition of culture that forces us to recognize that science and art are parallel aspects of the same basic creative impulse rather than unrelated pursuits. It is revealing to examine how artists and art museums are using the Internet; they are particularly sensitive to aesthetic issues, and eventually the ability to express meaning visually will help determine whether the Web succeeds or fails as a new medium.

The on-line experience currently has some serious limitations. Anyone who has explored the Web will have no trouble thinking of the effort as trying to beat the clock. Depending on the capabilities of your computer, modem and network server, you will soon become acquainted with the inadequate pace of information delivery, especially for pictures, movies or sound.

This condition is symptomatic not only of a medium in its infancy but also of a profound shift in the perception of ideas. A picture is literally worth a thousand words in the digital currency of bits and seconds. A page of text traveling via the Internet might arrive nearly instantaneously. A small, "thumbnail" image of the Mona Lisa might take a few seconds to reach your screen from the Louvre in Paris. Anything approaching the detail of the original painting might take months—and no video monitor could begin to display such detail. The computer has made obvious an understanding of the physical art object—what Marcel Duchamp once called a "delay"—a thing so laden with emotion and thought that it is capable of slowing our perception of time. The time of the on-line art object is very slow.

As a result, there is a distinct difference between an art book or gallery and an on-line source. In the physical world, it may take a while to find a book or get to a gallery, but once you are in the right place all the pictures are accessible immediately. The Web can take you "anywhere" instantly, but seeing the pictures there takes almost forever.

This situation makes for a curious tension with the Internet's role as a "store and forward" medium, one that allows the sender of a message to spend a great deal more time composing it than the receiver does reading it. This condition is very much like that of a painter who labors for months to produce an exhibition. At the opening of the show, the work is displayed as if it had happened all at once. The paintings have been "stored" by the painter until the gallery decides to "forward" them to the public. Paintings are stored information every bit as much as any other kind of data that one can download. The painful crawl at which digital art must arrive, however, highlights the fact that art on the Internet is not a thing, it is a condition: a conceptual place for an aesthetic idea to unfold and be communicated.

Scientists are excited by the prospect of placing research on the Internet, thereby increasing access to valuable databases. In a like way, the digital artist can broaden the composition and dissemination of an artistic condition, enticing the viewer into the process. A number of Web sites are not catalogues of preexisting works but rather aesthetic ex-

periences—anything from traditional sketchbooks, raw and uncensored, to attempts to use the Web as an interactive kinetic canvas, with which the person visiting the site participates in the creation of an artwork. What happens to the art if the computer crashes? Is this the ultimate ephemeral art? It is not just an artist's concern: in a medium where anyone can publish to everyone, the intellectual stability of peer-reviewed science journals sits cheek by jowl with the speculation of dilettantes.

Andy Warhol would obviously have loved the Web. What better way to play with the sacred images of high art than in a digital form? The Andy Warhol Museum would seem a natural starting point for our exploration. But, sadly, the Warhol museum site has none of the playfulness of the artist. Instead it simply shows views of the galleries and lists the paintings and prints in the museum's collection. Distinctly un-Warhol-like.

Museum sites in general are problematic: they function somewhat like exhibition catalogues, offering static records that may be interesting to the serious scholar but are often dull for the casual visitor. They also lack the physical presence of museum architecture that can instill in the viewer a contemplative experience. Some sites, such as the Whitney Museum or the Art Gallery of Ontario, are more innovative and attempt to compensate for this weakness in part by trying to get the Internet visitor engaged enough to send e-mail or even become an "electronic member" of the museum.

An odd hybrid site that mixes curatorial passion with a nonacademic approach is Roy Williams's Paintings of Vermeer. Williams, a researcher at the Center for Advanced Computing Research of the California Institute of Technology, is not a professional artist or museum employee. He is simply a guy who loves the work of Jan Vermeer (1632–1675). His Paintings of Vermeer site helps you "find your closest Vermeer" by locating all the paintings on a world map. You can also rank the paintings

THE CD EXAMINED

Everything Weather
The Weather Channel, 1996 (CD-ROM for Windows, $40)

Like the movie *Twister,* this CD-ROM blends nice visuals with a lackadaisical structure. Its oddest failure is its awkward linking of word and image, a key feature of multimedia. The disc does contain some excellent photographs (*such as the tornado above*), and the related text is informative, if a bit dry. The interactive features, however, are mostly shallow. Parents beware: the on-line access to weather forecasts costs $0.75 per call. —*Corey S. Powell*

BRIEFLY NOTED

FEYNMAN'S LOST LECTURE: THE MOTION OF PLANETS AROUND THE SUN, *by David L. Goodstein and Judith R. Goodstein. W. W. Norton, 1996 (book and audio CD, $35).*
Richard Feynman's prowess as an educator is well documented in this recently unearthed "lost" installment from his introductory physics lectures. With bulldog tenacity, he methodically reconstructs Isaac Newton's singular insight that the elliptical orbits of the planets can be explained entirely by the basic laws of geometry and gravity. The authors' chapters add helpful scientific background and reminiscences about Feynman. The greatest treat, however, is listening to Feynman's New York–accented voice reciting the lecture on the accompanying compact disc.

SUMMER STARGAZING: A PRACTICAL GUIDE FOR RECREATIONAL ASTRONOMERS, *by Terence Dickinson. Firefly Books, 1996 ($18.95).*
In this large-format volume, Terence Dickinson eschews star charts in favor of real photographs of the night sky. The result is an unusually attractive and straightforward introduction to the heavens (though urban dwellers might be disoriented by the many faint stars they will not be able to see). The book also includes a series of intelligent essays and a brief overview of the sky in the cooler seasons.

MINUTES OF THE LEAD PENCIL CLUB, *edited by Bill Henderson. Pushcart Press, 1996 ($22).*
The 40 contributors to this little volume share a deep skepticism about the value of the computers and Internet connections that increasingly permeate our society. There are some incisive moments here, but the overall dour tone does little to advance either the stabs at irony or the attempts at social commentary. And one has to wonder about the book's cluttered mixture of short essays, quotes and cartoons. Is it an alternative to hypertext or an unwitting homage?

Continued on page 87

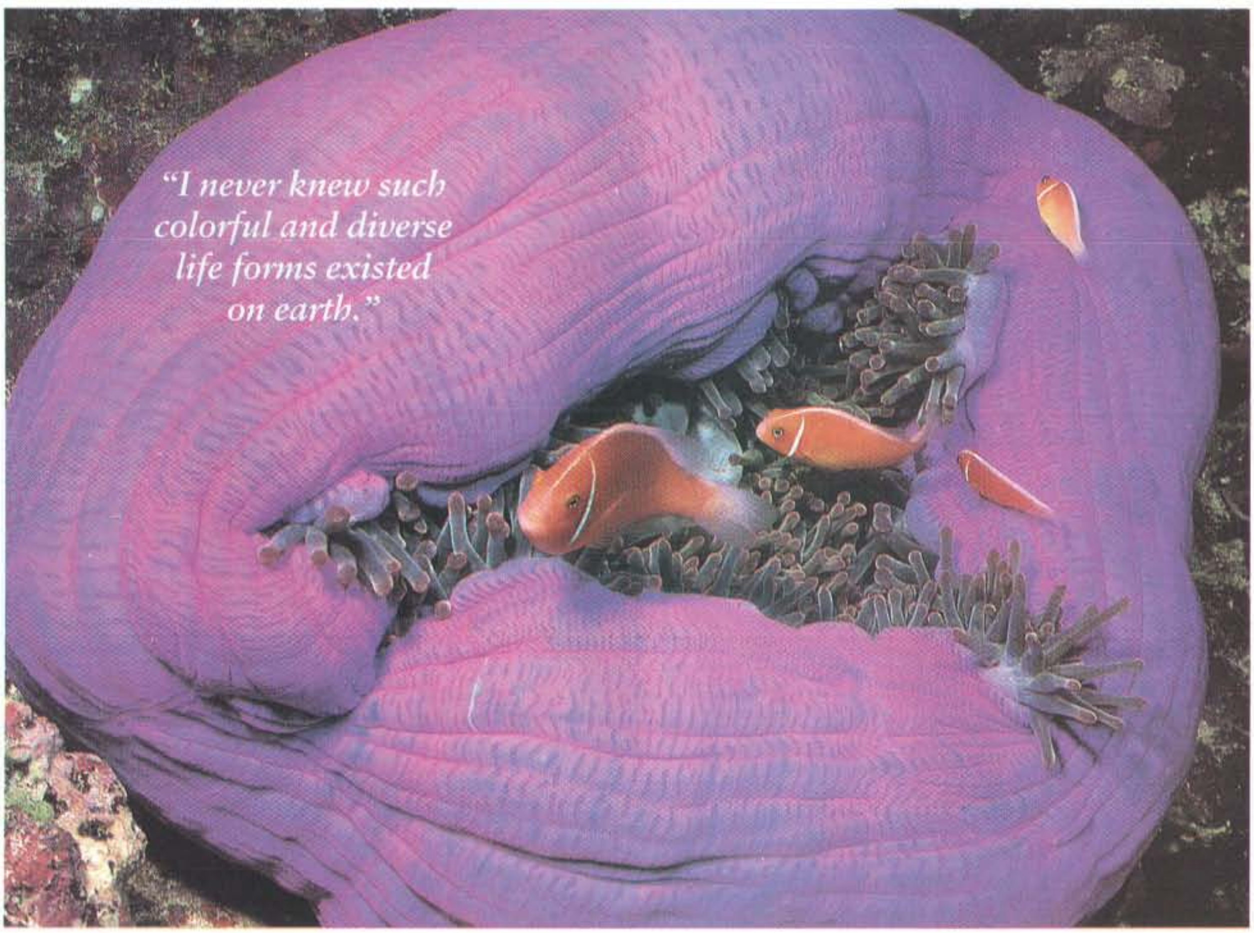
by popularity, based on how many "hits" each virtual painting on the Web site has received over an 11-day period. In a nod to a historical approach, Williams allows the user to view the paintings chronologically. He also includes his own interpretations of the paintings in an accompanying text.

How—or whether—Williams obtained permission to make digital copies of Vermeer's works is not revealed on his Web site. This missing information reminds us that the Web is a popular medium and that, in its present, unregulated form, the expert and the amateur exist side by side, for better and for worse. The enthusiasm engendered by the ease of access to the Web elevates the fan of Vermeer to digital curator status. Although professionals may rage against such disregard for law and convention, the Vermeer site boldly says that individuals may not know anything about art, but they know what they like, and they are not afraid to tell you about it.

The Warhol site and the Vermeer site, however, invite a more serious question about the nature of culture on the Web: What if I actually want to study at an academic level? Where do I go if I really am interested in detail? I joined the Getty in part to aid in answering such questions. The Getty Information Institute has made an effort to answer that question with a.k.a., an automated guide to cultural data.

The a.k.a. Web page contains a thesaurus of art-history terminology and a database of artists' and architects' names. Both can be used to guide searches through other archives, among them Getty's comprehensive index of current

THE ILLUSTRATED PAGE



Wonders of the Reef: Diving with a Camera
BY STEPHEN FRINK
Harry N. Abrams, 1996 ($39.95)

Stephen Frink is a professional photographer and instructor at Nikon's photography workshops. His writing is uneven: sometimes poetic, often full of technical details primarily of interest to fellow photographers. But there is no arguing with the results of Frink's innovative camera and lighting techniques. These underwater scenes shimmer with seemingly impossible colors—witness the clown fish and purple anemone (*above*)—and capture some remarkable moments, as when a member of the Underwater Explorers Society calmly feeds a Caribbean reef shark. —*Corey S. Powell*

literature on the history of Western art and an extensive database of architectural journals and magazines. Ultimately, the same basic suite of tools could make cultural databases all over the world available to anyone with a Web browser.

Elsewhere on the Getty Web site you will find another Getty print publication, "Introduction to Imaging: Issues in Constructing an Image Database." Writing about digital imagery is, after all, a task most appropriately carried out on the Web. This tutorial on digital art on-line has the "real" digital images that the print version can only represent in ink.

One of the most intriguing aspects of the Web, of course, is its ability to do more than passively gather information; it also allows the user to share in the process of what happens to that information. Two artists who have taken advantage of the interactivity and democracy of the Web are Komar & Melamid, Russian émigrés supported by the Dia Center for the Arts, a multidisciplinary contemporary arts organization based in New York City. They have approached the Web with irony and humor that strikes at one of the fundamentals of high art, "taste." Marshall McLuhan once called taste the first refuge of the insecure, and certainly Komar & Melamid have taken this thought to heart. Their site, The Most Wanted Paintings, is a homage to the notion that, by asking the right questions, anyone can create art.

Michael Govan, the director of Dia, explains: "The Most Wanted Paintings, as well as The Least Wanted Paintings, reflects the artists' interpretation of a professional market research survey about aesthetic preferences and taste in painting." Intending to discover what a true "people's art" would look like, Komar & Melamid request that the Web user fill out a survey that asks everything from "How carefully do you consider most of your spending decisions?" to "In general, would you rather see paintings of outdoor scenes, or would you rather see paintings of indoor scenes?" The artists then tabulate the results and create artworks that reflect the surveys.

Komar & Melamid, like many businesses, are interested in the Internet as a market research tool. The underlying questions, Govan explains, are those that many of us think of but few dare ask aloud: "What would art look like if it were to please the greatest number of people?" Or conversely: "What kind of culture is produced by a society that lives and governs itself by opinion polls?"

As a public art medium, the Web has a facility like no other public cultural architecture: not only can you talk back to it, you can even talk to other people who talk back to it. Consider Antonio Muntadas's File Room. Muntadas, a media artist, built a real, three-dimensional installation of file cabinets and documents about censorship and then created an evolving virtual file room on the Web. The project, hosted by the Randolph Street Gallery and installed in the Chicago Cultural Center, indexes cases of governmentally suppressed speech from classical Greek drama to contemporary art and journalism. An Internet user can search documents by country, date, medium censored or the kinds of views that have been silenced. Those who have been censored can add their own documents to the collection.

As the site's disclaimer asserts, "The File Room claims no scholarly, editorial or scientific authority, but instead proposes alternative methods for information collection, processing and distribution, to stimulate dialogue and debate around issues of censorship and archiving." Such debate is particularly apt in light of ongoing attempts to bar from the Internet itself material that some consider offensive or indecent.

Electronic communication has been a key factor in popular uprisings around the world during recent years (the role of the fax machine in publicizing China's Tiananmen demonstrations being perhaps the most notable instance). And the Web appears to be continuing that tradition by serving as a venue for art that might otherwise have no legitimate home. Art Crimes, for instance, is a gallery of graffiti art from a number of U.S. cities and 19 countries around the world. The site states, "Protect your history by making it digital. Tell your story, express your opinions and publish them. Make your own Web site and link up with us. Tell the world about your zine or your art enterprise.... Hook up."

Graffiti are almost always instantly painted over by other artists, gangs or by the authorities, so Art Crimes acts as an electronic archive of otherwise ephemeral works. The site also has articles, interviews, music and even a section on

BRIEFLY NOTED

Continued from page 85

SEARCHING FOR MEMORY: THE BRAIN, THE MIND, AND THE PAST, *by Daniel L. Schacter. Basic Books, 1996 ($27).*
Daniel Schacter has synthesized a broad overview (77 pages of footnotes and references!) of the ways that our brains store and, all too often, distort the past. Loss of memory is always unsettling; as illustrated by the legal dramas surrounding controversial "recovered memories" of abuse, the mind's blurring of objective truth also has the potential to ruin lives. Schacter argues passionately for a better appreciation of both the power and the limitations of our mental records.

THE ULTIMATE EINSTEIN. *Byron Preiss Multimedia, 1995 (CD-ROM for Windows or Macintosh, $49.95).*
The gorgeous, convoluted design of this CD-ROM offers yet another instance of multimedia smarts run amok. Serious and often thoughtful information is filtered through a needlessly complicated interface. Albert Einstein's ruminations appear on a fancy, rotating scroll that displays but 10 lines at a time; meanwhile concepts such as wave-particle duality, which would benefit from graphic treatment, are relegated to terse glossary definitions.

KRAKATAU: THE DESTRUCTION AND REASSEMBLY OF AN ISLAND ECOSYSTEM, *by Ian Thornton. Harvard University Press, 1996 ($39.95).*
What if you could wipe the earth clean and start all over again? Nature answered that question on a small scale in 1883, when the eruption of Krakatau, a volcano in Indonesia, obliterated essentially all the local flora and fauna. Within a generation the islands were again green, and today they teem with life. Ian Thornton's jargon-heavy style sometimes drowns out the inherent drama of the story, but the underlying message of life's remarkable resilience emerges loud and clear.

subway graffiti, including tips on how not to get hurt. This kind of site suggests that new art forms are part physical and part virtual. Can new kinds of scientific scholarship—possibly constructive, possibly corrosively inane—be far behind?

Other sites raise even more profoundly the question of just what is or is not art. By signing an Executive Order on February 23, 1995, directing the declassification of intelligence imagery acquired by the first generation of U.S. photoreconnaissance satellites, President Bill Clinton set the stage for a historical and geographic archive—and, in some sense, a marvelous trove of found art. In the opening decade of the 20th century, artists such as the young Pablo Picasso were inspired by the first aerial photographs, which revealed a startling way of seeing the landscape. Abstract forms and textures in these images appeared to be flat and dimensional at the same time, as if you could see a three-dimensional object (a cube, say) from all sides at once. What are these digital satellite photographs to the artists of today? On the Web, are they not common resource material for the artist as well as the scientist?

One of the great values of the Web is that it makes us take a fresh look at familiar concepts, inspiring a reexamination of rarely reconsidered categories like "art" and "science." Art is not only, as we commonly assume, about personalities, art objects, criticism, scholarship or history. It is about a deeper, slower reality that changes the way we understand the world. The Web presents art as a category of interest, but once we begin to investigate that category we find ourselves immersed in such a variety of interlinked territories that we must admit that the category is more than the sum of its parts.

The shift from memory to written record occurred some 5,000 years ago, when written property deeds replaced physical objects such as rings and daggers as seals of commitment. Today it seems we are moving from written record to collective electronic memory. This collective memory is open and accessible and presents us with a dynamic palette; it forces creativity on us. In a sense, looking for art on the Web *is* art on the Web, much like the gesture of a painter searching a canvas for a vision.

BEN DAVIS is program manager for communications at the Getty Information Institute in Santa Monica, Calif.

World Wide Web site addresses mentioned in the article:

The Andy Warhol Museum: http://www.clpgh.org/warhol/

The Whitney Museum of American Art: http://www.echonyc.com/~whitney/

The Art Gallery of Ontario: http://www.ago.on.ca/

Paintings of Vermeer: http://www.ccsf.caltech.edu/~roy/vermeer/index.html

The Getty Information Institute: http://www.ahip.getty.edu/

a.k.a. (Also Known As): http://www.ahip.getty.edu/aka/

Introduction to Imaging: Issues in Constructing an Image Database: http://www.ahip.getty.edu/intro_imaging/home.html

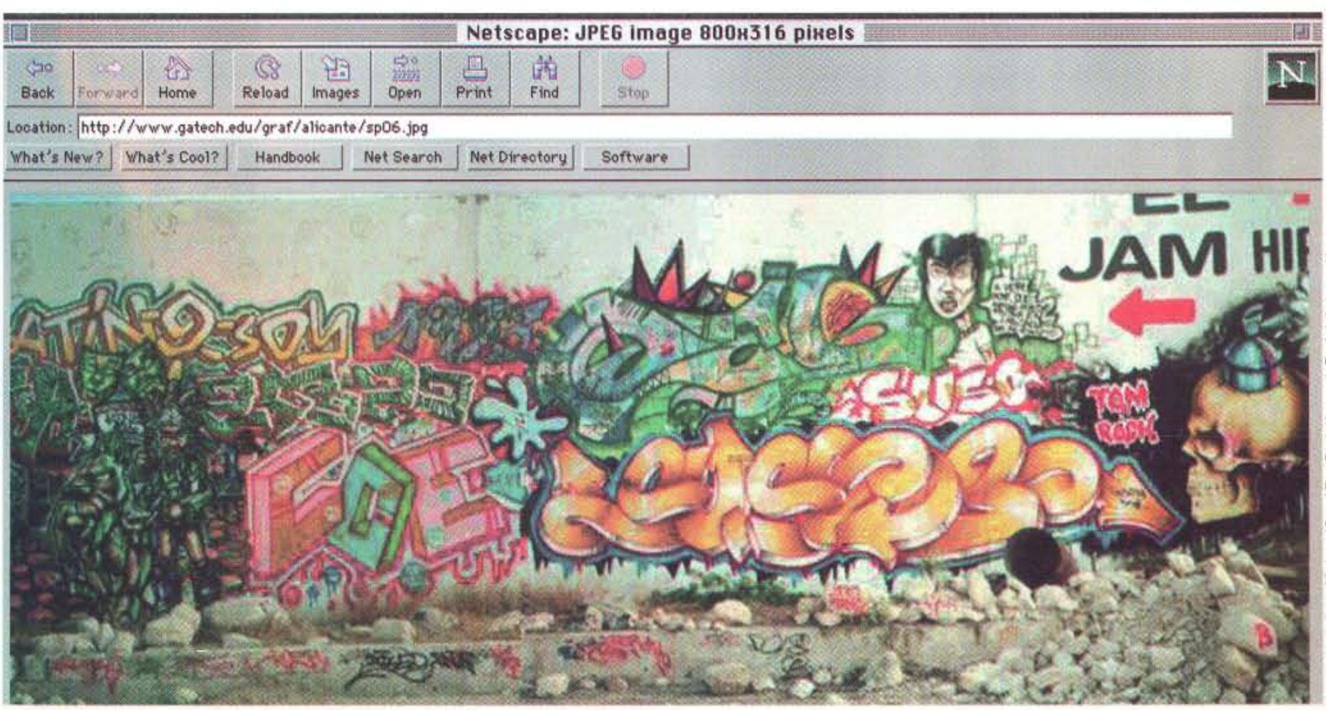
Dia Center for the Arts: http://www.diacenter.org/

The Most Wanted Paintings: http://www.diacenter.org/km/index.html

Antonio Muntadas's File Room: http://fileroom.aaup.uic.edu/FileRoom/documents/homepage.good.html

Art Crimes: http://www.gatech.edu/graf/

U.S. reconnaissance satellite images: http://edcwww.cr.usgs.gov/dclass/dclass.html



Fase2, Coe3, Susa.33, Loco13 and Tom Rock (*images*); Zock (*photograph*)

COMMENTARY

WONDERS

by Phylis and Philip Morrison

Great Books and City Crops

Cambridge, Massachusetts, is a long-urbanized, low-rise city of under 100,000, its population nearly constant for a century, with scant parkland or undeveloped lots. Gardens and lawns it has in plenty, but no one could point to a commercial farm. In 1995 one very successful local enterprise made surprising news, its agricultural technology at the highest level, its urban standing crop worth six figures in one indoor loft! Its 1/200th acre of deep-green marijuana beds were of a strain pampered for generous yield of illicit resin, won (and then busted) under a phalanx of fluorescent tubes glowing behind closed doors and blanked windows.

Twenty years before this covert, postmodern agronomy, we had grown real grain crops in our own Cambridge garden open to New England rain and sun. We sowed only a few square feet of wheat, for our minute city harvest was intended not as an economic act but as an educational one. By sharing the old complex experiences of growing our daily bread, we hoped to grasp the metaphors brought so long ago to poetry and language out of the fields of grain.

We began as novices. (Phil had never lived on a farm, although Phylis, a keen gardener, had spent childhood summers in Alpine fields.) One thing was plain: first get the seed. A city farmer seeks the phone book. The White Pages reach out 10 miles from downtown Boston to list what is still the nearest farm supplier. The friendly man who awaited our naive request did not understand at once what we sought. "Wheat seed?" he repeated, a bit puzzled. Of course: *all* wheat is seed. What we wanted is always called seed wheat. He brought us the smallest quantity he could easily weigh out. It was a sight, dyed a daunting, inedible pink to announce its mercurial antifungal dressing. We would not be tempted to eat these seeds before planting.

Our Cambridge backyard is not large: we planted a small first plot with the seed. Soon enough, over the summer weeks, a few dozen pale-green shoots sprouted to 16 or 20 inches high. But that was all they did, although they had seemed such healthy blades of grass. Autumn came, then winter, and every shoot died back. We had failed utterly in our amateur farming.

One day as winter wore on, Phil called in excitement from the downstairs room, where many books arrive each day, their senders hopeful of a review. "It's *winter* wheat!" A volume on wheat and its characteristics had arrived (written by R. F. Peterson), and he had quickly snatched the point. We were agronomically on schedule through the time of first frosts. We were painstakingly assured in print that once spring came, the blades would lengthen and green again. The vital grass would blossom and form a head; by full summer the seeds would fill and ripen. That is the behavior of the most important North American varieties of wheat.

So it came to pass. In mid-July we found a sickle to harvest the ripened grain. Then we threshed it with our fingers, piling grains of wheat along with the glumes and light chaff. We winnowed the threshed grain from the rest by dropping it into a bowl while we fanned hard to blow away the residues. Then we ground our handful of grain by mortar and pestle into coarse flour.

We were assured in print that once spring came, the blades would lengthen and green again.

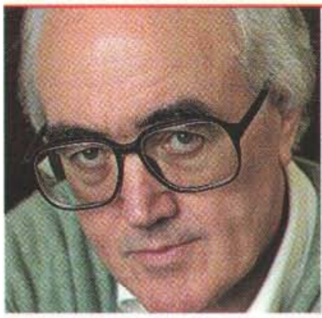
VLAD GUZNER

That year our Thanksgiving dinner was graced by a rather heavy roll, made ceremonially from our first crop.

We had so clear a sense of learning that we tried again the next year on a larger scale, a weeded "field" now a couple of yards on a side. By then neighborhood health food stores had arrived, and it was easy to buy wheat we could choose either to eat or to sow. Soon we learned why the most successful wheats of the Green Revolution (and the rice, too) had been bred to inherit the short habit of dwarfed varieties. When the wheat was nearing maturity, a fierce July thunderstorm knocked most of our top-heavy plants over every which way. We faced the farmers' bane: lodged grain, heads beaten flat to the earth. Harvesting that tangled and muddied crop from the ground was difficult; we saved all we reasonably could.

Some days later our neighbor's daughters came to our garden door. The two wondered if they could still gather some good of our tumbled wheat. The young women gleaners in our field evoked that Ruth who bore King David's grandfather

Continued on page 91



CONNECTIONS
by James Burke

A Lot of Boloney

I have to confess a fatal weakness for Bologna, Italy. Apart from having one of the oldest universities in Europe, and possibly the most elegant women on the planet, it just so happens to be the food capital of the known universe. And after lunching on *tortellini alla panna,* you can go savor another work of mouthwatering precision: a giant brass meridian line, inlaid across the floor of the city's cathedral. Put there by Gian Domenico Cassini, the hottest astronomer around in 1668. Which was when his reputation brought him an offer he couldn't refuse, from Louis XIV's right-hand man, Jean-Baptiste Colbert: to run the new Paris observatory. Subsequently, he joined the great national effort there to determine the shape of the earth (which the French thought was not flattened at the poles).

Businessmen instantly got stuck on this exciting idea of pushing the communications envelope.

Colbert needed to know such arcana so that the new navy he was putting together would more accurately be able to relate star-fix angles to positions on the planetary high seas (which would be different on an earth that was, or was *not,* flattened at the poles). This way, French ships would be able to navigate better. And rule the waves. And, perhaps, give the English one in the eye, by snitching the prime meridian from Greenwich and moving it to Paris. Unfortunately for *amour-propre,* they were wrong about the shape of the globe, which is why I'm writing this in Greenwich mean time.

The sidereal shenanigans were key to Colbert's grand plan to make France a mercantile superpower, as part of which Colbert also offered tax breaks to anyone interested in sailing (more accurately, it was now hoped) to exotic lands and returning with import deals for high-end consumables. Idea being, this trade could become a French monopoly and make oodles of ecus for king and country. Well, king.

This perfectly legitimate scheme for avoiding taxation was yet another offer too good to refuse, so in no time at all freebooters were hauling back shiploads of gold, ivory, slaves and gum from Senegal in West Africa. Senegal gum turned out to be just what you needed to machine-print chintz (the newest Euro-fashion craze from India) with fast colors, because the gum acts as a dye binding agent. By mid-18th century the chintziest guy in town was an Irishman called Francis Nixon, who had a method for producing all the cheap prints you wanted, and at high speed. Nixon's trick involved pressing steel designs onto copper rollers, then coloring up the rollers and running cotton between them. Thus inventing matching curtains and covers.

Anyway, by 1818 Jacob Perkins of Newburyport, Mass., had added a few modifications to Nixon's process and went to London to persuade the Bank of England to give him the printing contract for banknotes whose intricate designs he claimed nobody could copy. Speedily—18 years later—the bank said yes. (Well, this *was* England.) Four years later the patient Perkins really licked his competitors when he landed the job of producing the "Penny Black," the world's first postage stamp.

Businessmen all over instantly got stuck on this exciting idea of pushing the communications envelope. By 1874, thanks to the introduction of the stamp to Switzerland, Bern was home to the Universal Postal Union, and the international community was dividing mail into three categories: letters, parcels and a brand-new thing called a postcard.

VLAD GUZNER

Illustrations on some of the earliest cards came from the pen of Phil May, a British cartoonist who did his best work for an up-and-coming satirical magazine, *Punch*. The magazine hadn't planned to include drawings, but events presented one of those opportunities every editor dreams of. When Queen Victoria's husband, Prince Albert, came up with a competition to decide who should fresco the interiors of the newly-rebuilt-in-19th-century-imitation-Gothic Houses of Parliament, the entries His Royal Highness liked best were so ludicrously *dreadful,* there was only one way to stop him. Publish them. It worked.

But not even *Punch* could prevent the Gothic revival. Next time you're in Britain, note the many 19th-century churches, originally built with gargoyles and gaslight. Gothic was cheaper than neoclassical, so the Victorian church commissioners threw up more than 500 of them. The fault, I suppose, of that back-to-the-Middle-Ages lunacy we now call Romanticism, spearheaded (like that medieval touch?) by a young German philosopher—Johann G. Herder—who was deeply into the fundamental unity of humankind, nature, German folk songs and the stuff of Sturm und Drang—an epic view of existence, but, by our standards...well over the top!

Herder had been attuned to such excesses by the clamorous arrival in Germany in the late 18th century of a collection of Gaelic poems written by the

third-century Irish warrior-poet Ossian. These took Europe and Herder by sturm—sorry, storm. For Romantics, the works pulsated with the pure and powerful feelings of a primitive people. And triggered the Romantic movement. Of such things are great moments of history made.

Too bad the poems were fake. "Discovered" by a middle-rate Scots poet named James Macpherson, who bundled a few ballads he'd collected (on a tour of Scotland) together with his own work and passed them off as Gaelic translations of 1,500-year-old originals. Still, he did help give us Romanticism, which brought us pathology and radio (more of that in another column). Why would somebody like Macpherson be on an antiquarian ballad hunt in the first place? Perhaps because back then the future of Scottish culture was looking bleak.

Ever since 1715, the Catholic Stuarts had been submitting claims to the English throne (now occupied by Protestant Germans) in the form of armed uprisings. Redcoats were all over the Highlands. And matters such as clans, tartans and speaking the local lingo were all stuff that could get you seriously hanged. The English even wrote a special extra verse to their national anthem, all about "rebellious Scots to crush!"

Things came to a head in 1745, when the last royal Stuart, Bonnie Prince Charlie, and his "murderous band of cutthroats" (a.k.a. "band of brave patriots") got as far south as Derby, which caused a run on the pound. Now, you don't mess with the Bank of England and get away with it. But get away he did. In the words of the song: "Over the Sea to Skye," and then across the channel to the Continent. To this day, in memory of Charlie's flight to foreign parts, romantic Scots will raise their glasses to "the king over the water."

And finally: the reason I began this particular example of tortuously crafted boloney the way I did. Because guess where Charlie ended up spending the best of his declining years in exile? Where would you go, if (like him) you were looking for intellectual chat, the company of elegant women and overindulgence in food and drink (of which the prince ultimately expired)? Whichever way you sliced it, there was only one choice: Bologna. SA

Wonders, continued from page 89

after asking, "Let me glean . . . among the sheaves."

Now the precise old words of art were made vivid. First had we sown, and then tarried, to reap in haste after the storm laid flat the stalks of our tallish strain. Others who were landless gleaned a little of our lost bounty. Grinding to coarse flour was hard work still, if not so bitter as the toil of the old serving woman who ground meal daily for that arrogant crowd of suitors lounging in Ithaca, eager to claim Penelope and her legacy from royal Odysseus. The glorious language in the Scriptures and in Homer was at least a little our own by right.

The domestication of grains spans 11,000 years (with clear precursors long before), but the next century is likely to see unprecedented change. The Green Revolution was apt opportunism: the farmers' grief over lodged grain had stimulated breeders to work on crosses with dwarf strains for both wheat and rice. Stiff, short straws supported bigger grainheads without buckling. More important, yield increased under large additions of fertilizer.

It won't be so simple for the next vital task, doubling our crops perhaps for the last time. Much detailed attention to the spectrum of crop limitations in the diverse hands of farmers is needed. Ahead are 50 years of modest but locally effective novelties in field management of the crop cycle and in new, precisely tailored strains. Other major crops—say, improved millets and sorghums for drylands or even leguminous nodules bioengineered into the cereal roots to replace nitrate fertilizer—may well appear.

After a full century of high biotechnology, an optimist may conjecture that a stabilizing world of 10 or 12 billion people will have outdone today's best marijuana lofts. Many farmers might culture only plant tissues or cells or tiny organelles under biopolymer plastic. Scarce inputs of water and nutrients can be conserved and recycled even in the desert. Then the harvest index can approach its ultimate limit, most products taken directly as sugars, starches and proteins, supplemented to suit by colors, flavors and textures. No carefree Eden, it will be a planetary garden worked with powerful, growing knowledge and yet by much sweat of the brow. We can be well but frugally fed, with little waste. SA

WORKING KNOWLEDGE

SLOT MACHINES

by Raymond Heidel

Casino revenues in the U.S. during 1994 totaled \$17.5 billion. Of that, 65 percent was funneled as nickels, quarters and dollar bills through slot machines. These devices may qualify as the world's most user-friendly computers. The collection of springs, gears, levers and weights that the inventor Charles Fey cobbled together at the turn of the century to make the Liberty Bell, which became the model for the familiar three-reel slot machine, has increasingly given way to a conglomeration of microprocessors and memory chips.

Each reel can stop at 22 positions (each one displaying various symbols, such as fruit, or simply a blank space), for a total of 10,648 different combinations. When slot machines were purely mechanical, the maximum odds of a payout were therefore 10,648 to one. But the heart of a contemporary slot machine is a microprocessor programmed to generate random numbers, which can be assigned to any combination on the reels. In effect, the microprocessor dictates what the machine will display—and pay. Because many or few random numbers can be assigned to any given combination, makers of slot machines can vary the odds as desired. For the top jackpot, the odds of pulling the single right combination may be set at 10 million to one.

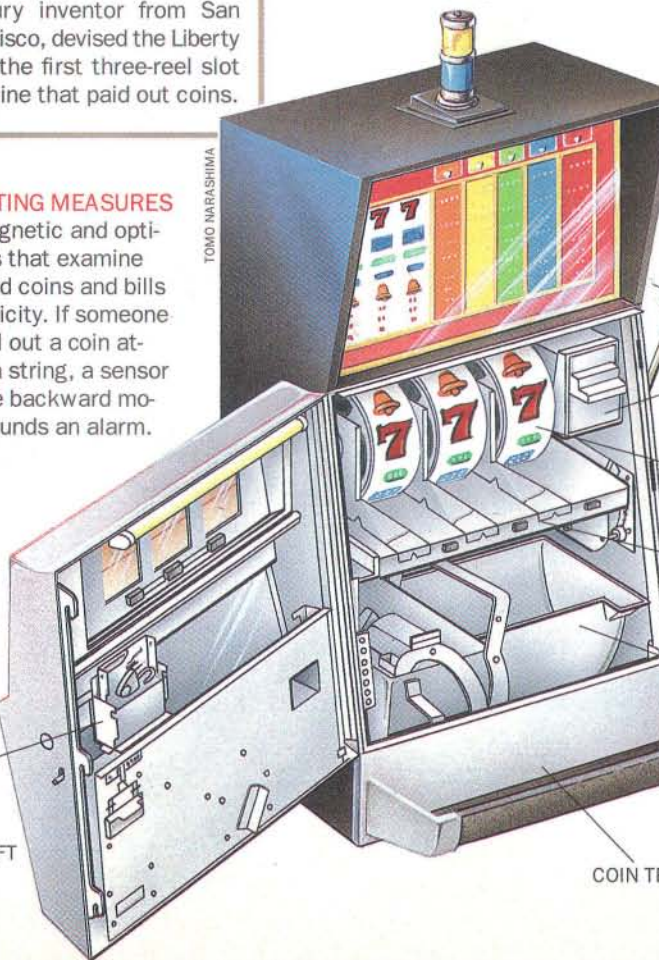
Offering higher odds has meant that jackpots in the 23 states with casino gambling can balloon into the millions of dollars. And as the stakes have mushroomed, more people have flocked to play the one-armed bandits.

RAYMOND HEIDEL is vice president of engineering for Bally Gaming in Las Vegas, Nev.

LIBERTY BELLE BOOKS

CHARLES FEY, a turn-of-the-century inventor from San Francisco, devised the Liberty Bell, the first three-reel slot machine that paid out coins.

MICROPROCESSOR issues a command to start the spinning of the three reels. (In a video slot machine, a graphic image simulates the motion of reels.) It also produces random numbers that correspond to combinations of symbols on the reels—for instance, cherries, bars or "jackpots." The hopper releases a jangle of coins after a winning hit.



TOMO NARASHIMA

ANTICHEATING MEASURES include magnetic and optical sensors that examine the inserted coins and bills for authenticity. If someone tries to pull out a coin attached to a string, a sensor detects the backward motion and sounds an alarm.

LAURIE GRACE

NETWORKING over telephone lines in Nevada coordinates the odds and payouts of hundreds of slot machines across the state. Each time a machine is played, the potential winnings of its jackpot rise. Last October this type of progressive game yielded a payoff of nearly \$11 million, the largest ever.